

#5
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the application of:)
Hirofumi, ITO, Kozo, AKIYOSHI, and)
Nobuo, AKIYOSHI)
Serial No.:)
Filed:)
For: IMAGE GENERATING METHOD,)
APPARATUS AND SYSTEM)
USING CRITICAL POINTS)

Group Art Unit:

Attorney Docket:

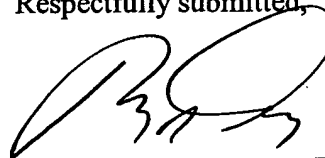
Commissioner of Patents and Trademarks
Washington, D.C. 20231
U.S.A.

Dear Sir/Madam:

LETTER ENCLOSING CERTIFIED COPY
OF PRIORITY DOCUMENT

To complete the claim for foreign priority made under 35 U.S.C. § 119(a)-(d) or (f) in
respect of the above-noted application, the applicant encloses the certified copy of Japanese
Patent Application No. 2001-29061, dated February 6, 2001.

Respectfully submitted,



Ralph A. Dowell
Registration No. 26,868

DOWELL & DOWELL P.C.
Suite 309
1215 Jefferson Davis Highway
Arlington, Virginia
U.S.A. 22202
Tel: (703) 415-2555
Fax: (703) 415-2559

Date: 1/28/2002
Enclosures
92637-17/NWH/mis/January 25, 2002



PATENT OFFICE

JAPANESE GOVERNMENT

Jc672 U.S. PTO
10/056025
01/28/02

This is to certify that the annexed is a true copy of the following application as filed with this Office.

DATE OF APPLICATION : February 6, 2001

APPLICATION NUMBER : Patent Application 2001-29061

APPLICANT(S) : MONOLITH CO., LTD.

April 27, 2001

Commissioner, Patent Office

Kozo OIKAWA

日 本 国 特 許 庁
JAPAN PATENT OFFICE

Jc872 U.S. PRO
10/056025
01/28/02

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年 2月 6日

出 願 番 号

Application Number:

特願2001-029061

出 願 人

Applicant(s):

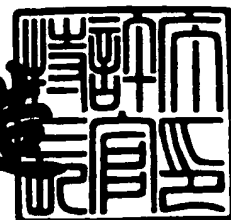
株式会社モノリス

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 4月27日

特許庁長官
Commissioner,
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3035406

【書類名】 特許願

【整理番号】 MN-0129

【提出日】 平成13年 2月 6日

【あて先】 特許庁長官殿

【国際特許分類】 H04B 14/00

【発明者】

 【住所又は居所】 東京都港区麻布十番 1 - 7 - 3 株式会社モノリス内

 【氏名】 伊藤 博文

【発明者】

 【住所又は居所】 東京都港区麻布十番 1 - 7 - 3 株式会社モノリス内

 【氏名】 秋吉 仰三

【発明者】

 【住所又は居所】 東京都港区麻布十番 1 - 7 - 3 株式会社モノリス内

 【氏名】 秋吉 信雄

【特許出願人】

 【識別番号】 396001980

 【氏名又は名称】 株式会社モノリス

【代理人】

 【識別番号】 100105924

 【弁理士】

 【氏名又は名称】 森下 賢樹

 【電話番号】 0422-23-7415

【手数料の表示】

 【予納台帳番号】 054977

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

特 2 0 0 1 - 0 2 9 0 6 1

【包括委任状番号】 9717729

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 画像生成方法、装置およびシステム

【特許請求の範囲】

【請求項 1】 キーフレームおよびそれらの間の対応点情報をもとに中間フレームを生成する方法において、

中間フレームを生成する装置の処理性能を特定する工程と、

特定された処理性能に応じて、生成すべき中間フレームの解像度を決定する工程と、

を含むことを特徴とする画像生成方法。

【請求項 2】 決定された解像度にしたがって補間計算により中間フレームを生成する工程をさらに含む請求項 1 に記載の方法。

【請求項 3】 前記処理性能を特定するために、前記装置におけるオプションなアクセラレータの有無が判定される請求項 1、2 のいずれかに記載の方法。

【請求項 4】 前記処理性能を特定するために、前記装置の CPU の処理速度が判定される請求項 1 から 3 のいずれかに記載の方法。

【請求項 5】 前記特定する工程が前記装置にて実施される請求項 1 から 4 のいずれかに記載の方法。

【請求項 6】 前記特定する工程は、前記装置にて実験的に中間フレームを生成させ、その状況をもとに処理性能を特定する請求項 1 から 5 のいずれかに記載の方法。

【請求項 7】 キーフレームおよびそれらの間の対応点情報をもとに中間フレームを生成する方法において、

中間フレームを生成する装置の処理性能を特定する工程と、

特定された処理性能によってリアルタイムの再生が可能な範囲の中間フレームの解像度を決定する工程と、

決定された解像度で中間フレームを生成する工程と、

を含むことを特徴とする画像生成方法。

【請求項 8】 前記決定する工程にて、低い解像度でもリアルタイムな生成

が困難と判定された場合、緩和された速度で中間フレームを生成する工程と、
生成された中間フレームをムービーファイルへ変換して記録する工程と、
をさらに含む請求項 7 に記載の方法。

【請求項 9】 前記特定する工程は、前記装置にて実験的に中間フレームを生成させ、その状況をもとに処理性能を特定する請求項 7、8 のいずれかに記載の方法。

【請求項 10】 キーフレームおよびそれらの間の対応点情報をもとに中間フレームを生成する装置において、

当該装置の処理性能を特定する性能特定部と、

特定された処理性能に応じて、生成すべき中間フレームの解像度を決定する解像度決定部と、

を含むことを特徴とする画像生成装置。

【請求項 11】 決定された解像度にしたがって補間計算により中間フレームを生成する中間フレーム生成部をさらに含む請求項 10 に記載の装置。

【請求項 12】 前記性能特定部は、処理性能を特定するために、当該装置におけるオプションなアクセラレータの有無を検出する請求項 10、11 のいずれかに記載の装置。

【請求項 13】 前記性能特定部は、処理性能を特定するために、当該装置の CPU の処理速度を判定する請求項 10 から 12 のいずれかに記載の装置。

【請求項 14】 前記性能特定部は、当該装置にて実験的に中間フレームを生成させ、その状況をもとに当該装置の処理性能を特定する請求項 10 から 13 のいずれかに記載の装置。

【請求項 15】 キーフレームおよびそれらの間の対応点情報をもとに中間フレームを生成する装置において、

中間フレームを生成する装置の処理性能を特定する性能特定部と、

特定された処理性能によってリアルタイムの再生が可能な中間フレームの解像度を決定する解像度決定部と、

決定された解像度で中間フレームを生成する中間フレーム生成部と、

を含むことを特徴とする画像生成装置。

【請求項 1 6】 前記解像度決定部は、低い解像度でもリアルタイム生成が困難と判定された場合、リアルタイム処理を断念すべくその旨を中間フレーム生成部へ通知し、

中間フレーム生成部は、緩和された速度で中間フレームを生成し、

当該装置はさらに、生成された中間フレームをムービーファイルへ変換して記録する変換部を含む請求項 1 5 に記載の装置。

【請求項 1 7】 前記性能特定部は、当該装置にて実験的に中間フレームを生成させ、その所要時間を計測して処理性能を特定する請求項 1 5、1 6 のいずれかに記載の装置。

【請求項 1 8】 前記性能特定部は、ユーザの指示に基づいて処理性能を定めることを特徴とする請求項 1 5 から 1 7 のいずれかに記載の装置。

【請求項 1 9】 前記キーフレームおよび対応点情報を外部より取得する通信部をさらに含む請求項 1 5 から 1 8 のいずれかに記載の装置。

【請求項 2 0】 サーバおよびクライアントを含むシステムにおいて、
サーバは、キーフレームおよびそれらの間の対応点情報を保持する保持部と、それらキーフレームおよび対応点情報をクライアントへ送信する通信部を含み、
クライアントは、送信されたキーフレームおよび対応点情報を取得する通信部と、取得したキーフレームおよび対応点情報を保持する保持部とを含み、
前記サーバまたはクライアントは、そのクライアントの処理性能を特定する性能特定部を含み、

前記クライアントはさらに、特定された処理性能によって所望の速度による再生が可能な中間フレームの解像度を決定する解像度決定部と、決定された解像度で中間フレームを生成する中間フレーム生成部とを含むことを特徴とする画像生成システム。

【請求項 2 1】 前記サーバは、キーフレームどうしのマッチングを、特異点を抽出することにより計算し、前記対応点情報を生成する請求項 2 0 に記載のシステム。

【請求項 2 2】 コンピュータにて実行可能なプログラムであって、このプログラムは、

キーフレームおよびそれらの間の対応点情報をもとに中間フレームを生成する装置の処理性能を特定する工程と、

特定された処理性能に応じて、生成すべき中間フレームの解像度を決定する工程と、

を実行せしめることを特徴とするコンピュータプログラム。

【請求項 2 3】 決定された解像度にしたがって中間フレームを生成する工程をさらに実行せしめる請求項 2 2 に記載のプログラム。

【請求項 2 4】 コンピュータにて実行可能なプログラムであって、このプログラムは、

キーフレームおよびそれらの間の対応点情報をもとに中間フレームを生成する装置の処理性能を特定する工程と、

特定された処理性能によってリアルタイムの再生が可能な範囲の中間フレームの解像度を決定する工程と、

決定された解像度で中間フレームを生成する工程と、

を実行せしめることを特徴とするコンピュータプログラム。

【請求項 2 5】 前記解像度を決定する工程にて、低い解像度でもリアルタイム生成が困難と判定された場合、緩和された速度で中間フレームを生成する工程と、

生成された中間フレームをムービーファイルへ変換して記録する工程と、

をさらに実行せしめる請求項 2 4 に記載のコンピュータプログラム。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

この発明は、画像生成技術に関し、とくに補間技術を利用して画像を生成する方法、装置、システムに関する。

【0 0 0 2】

【従来の技術】

B S デジタル放送が開始され、また多数のユーザが P C や携帯電話からインターネット上のサービスを享受するようになり、放送と通信の垣根が急速に崩れつ

つある。ブロードバンド時代においてマルチメディア文化は真の進展を遂げるであろうし、その中でもとりわけ動画の配信はキーテクノロジーとなる。人間が外界から取得する情報の量において、映像は音声の200倍の内容を伝達できるといわれる。映像は娯楽用途にとどまらず、広く人間の生活と文化を支えるソフトインフラとなる。映像がデジタル化されるにつれ、CGや画像処理技術の適用によって画像関連技術の活躍の場も広がる。

【0003】

【発明が解決しようとする課題】

たとえば、動画の圧縮符号化技術としてMPEG (Motion Picture Expert Group) が知られる。MPEGは、MPEG1、2、4と進化を遂げ、標準技術として今後もさらなる拡張が予定されている。しかしながら、いわゆるブロックノイズとしてその問題点が指摘され、また、さらなる圧縮率と高画質の両立についてはこの分野における永続的な課題である。

【0004】

この発明はそうした状況に鑑みてなされたものであり、その目的は、動画の生成を比較的軽いデータで実現する技術の提供にある。本発明の別の目的は、ユーザ個別に最適な動画の生成を実現する技術の提供にある。本発明のさらに別の目的は、動画を少ないデータから比較的高画質に生成する技術の提供にある。

【0005】

【課題を解決するための手段】

本発明のある態様は、画像生成方法に関する。この方法はキーフレームおよびそれらの間の対応点情報をもとに中間フレームを生成する方法であり、中間フレームを生成する装置の処理性能を特定する工程と、特定された処理性能に応じて、生成すべき中間フレームの解像度を決定する工程とを含む。また、決定された解像度にしたがって中間フレームを生成する工程をさらに含んでもよい。

【0006】

対応点情報は、本出願人が先に特許第2927350号にて提案した画像マッチング技術（以下「前提技術」という）によって生成されてもよいが、もちろんそれに限られない。

【 0 0 0 7 】

前記処理性能を特定するために、オプションなアクセラレータの有無、CPUの処理速度またはタイプなど、またはこれらの組合せを判定してもよいし、画像生成装置に実験的に中間フレームを生成させ、その状況、たとえば所要時間を計測してもよい。

【 0 0 0 8 】

本発明の別の態様も画像生成方法に関する。この方法は、中間フレームを生成する装置の処理性能を特定する工程と、特定された処理性能によってリアルタイムの再生が可能な範囲の中間フレームの解像度を決定する工程と、決定された解像度で中間フレームを生成する工程とを含む。また、前記決定する工程にて、低い解像度でもリアルタイム生成が困難と判定された場合、緩和された速度で中間フレームを生成する工程と、生成された中間フレームをムービーファイルへ変換して記録する工程とをさらに含んでもよい。

【 0 0 0 9 】

本発明のさらに別の態様は、画像生成装置に関する。この装置は、中間フレームを生成する装置の処理性能を特定する性能特定部と、特定された処理性能に応じて、生成すべき中間フレームの解像度を決定する解像度決定部とを含む。また、決定された解像度にしたがって中間フレームを生成する中間フレーム生成部をさらに含んでもよい。

【 0 0 1 0 】

本発明のさらに別の態様も画像生成装置に関する。この装置は、中間フレームを生成する装置の処理性能を特定する性能特定部と、特定された処理性能によってリアルタイムの再生が可能な中間フレームの解像度を決定する解像度決定部と、決定された解像度で中間フレームを生成する中間フレーム生成部とを含む。また、解像度決定部は、低い解像度でもリアルタイム生成が困難と判定された場合、リアルタイム処理を断念すべくその旨を中間フレーム生成部へ通知し、中間フレーム生成部は、緩和された速度で中間フレームを生成してもよく、その場合、当該装置はさらに、生成された中間フレームをムービーファイルへ変換して記録する変換部を含んでもよい。

【 0 0 1 1 】

性能特定部は、ユーザの指示に基づいて処理性能を定めてもよく、この場合、事実上ユーザが処理性能を決定してもよい。処理性能を指定することにより、ユーザが望む再生形態が選択できる意味をもつ。

【 0 0 1 2 】

本発明のさらに別の態様は、サーバクライアントシステムに関する。サーバは、キーフレームおよびそれらの間の対応点情報を保持する保持部と、それらキーフレームおよび対応点情報をクライアントへ送信する通信部を含む。クライアントは、送信されたキーフレームおよび対応点情報を取得する通信部と、取得したキーフレームおよび対応点情報を保持する保持部とを含む。また、サーバまたはクライアントは、そのクライアントの処理性能を特定する性能特定部を含み、クライアントはさらに、特定された処理性能によって所望の速度による再生が可能な中間フレームの解像度を決定する解像度決定部と、決定された解像度で中間フレームを生成する中間フレーム生成部とを含む。サーバは、例えば前提技術に基づき、キーフレームどうしのマッチングを、特異点を抽出することにより計算し、対応点情報を生成してもよい。

【 0 0 1 3 】

なお、以上の各構成、工程を任意に入れ替えたり、方法と装置の間で表現を一部または全部入れ替え、または追加したり、工程をコンピュータプログラムによる処理に変更したり、発明の形態を記録媒体等に変更したものもまた、本発明として有効である。

【 0 0 1 4 】

【発明の実施の形態】

はじめに、実施の形態で利用する多重解像度特異点フィルタ技術とそれを用いた画像マッチング処理を「前提技術」として詳述する。これらの技術は本出願人がすでに特許第 2 9 2 7 3 5 0 号を得ている技術であり、本発明との組合せに好適である。本発明では、画像上にメッシュを設け、その格子点によって多数の画素を代表されるため、もともと前提技術のような画素単位のマッチング技術に対する適用効果が高いためである。ただし、実施の形態で採用可能な画像マッチン

グ技術はこれに限られない。

図18以降、前提技術を利用した画像補間技術および映像効果技術を具体的に説明する。

【0015】

[前提技術の背景]

ふたつの画像の自動的なマッチング、つまり画像領域や画素どうしの対応付けは、コンピュータビジョンやコンピュータグラフィックスにおける最も難しくかつ重要なテーマのひとつである。例えば、あるオブジェクトに関して異なる視点からの画像間でマッチングがとれれば、他の視点からの画像を生成することができる。右目画像と左目画像のマッチングが計算できれば、立体画像を用いた写真測量も可能である。顔の画像のモデルと他の顔の画像のマッチングがとれたとき、目、鼻、口といった特徴的な顔の部分を抽出することができる。例えば人の顔と猫の顔の画像間でマッチングが正確にとられたとき、それらの中割画像を自動的に生成することでモーフィングを完全自動化することができる。

【0016】

しかし従来一般に、ふたつの画像間の対応点は人がいちいち指定しなければならず、多大な作業工数を要した。この問題を解消するために数多くの対応点自動検出方法が提案されている。例えば、エピポーラ直線を用いることによって対応点の候補の数を減らす考えがある。しかし、その場合でも処理はきわめて複雑である。複雑さを低減するために、左目画像の各点の座標は通常右目画像でもほぼ同じ位置にあると想定される。しかし、こうした制約を設けると、大域的特徴及び局所的特徴を同時に満たすマッチングをとることは非常に困難になる。

【0017】

ボリュームレンダリングでは、ボクセルを構成するために一連の断面画像が用いられる。この場合、従来一般に、上方の断面画像における画素が下方の断面画像の同一箇所にある画素と対応すると仮定され、これらの画素のペアが内挿計算に用いられる。このようにきわめて単純な方法を用いるため、連続する断面間の距離が遠く、オブジェクトの断面形状が大きく変化する場合、ボリュームレンダリングで構築されたオブジェクトは不明瞭になりがちである。

【0018】

立体写真測量法など、エッジの検出を利用するマッチングアルゴリズムも多い。しかしこの場合、結果的に得られる対応点の数が少ないため、マッチングのとれた対応点間のギャップを埋めるべく、ディスパリティの値を内挿計算しなければならない。一般にあらゆるエッジ検出器は、それらが用いる局所的なウィンドウの中で画素の輝度が変化したとき、これが本当にエッジの存在を示唆するかどうかを判断することが難しい。エッジ検出器は、本来的にすべてハイパスフィルタであり、エッジと同時にノイズも拾ってしまう。

【0019】

さらに別の手法として、オプティカルフローが知られている。二枚の画像が与えられたとき、オプティカルフローでは画像内のオブジェクト（剛体）の動きを検出する。その際、オブジェクトの各画素の輝度は変化しないと仮定する。オプティカルフローでは例えば (u, v) のベクトル場の滑らかさといった、いくつかの付加的な条件とともに、各画素の動きベクトル (u, v) を計算する。しかし、オプティカルフローでは画像間の大域的な対応関係を検出することはできない。画素の輝度の局所的な変化に注目するのみであり、画像の変位が大きい場合、システムの誤差は顕著になる。

【0020】

画像の大域的な構造を認識するために、多重解像度フィルタも数多く提案されてきた。それらは線形フィルタと非線形フィルタに分類される。前者の例としてウェーブレットがあるが、線形フィルタは一般に、画像マッチングにはさして有用ではない。なぜなら、極値をとる画素の輝度に関する情報がそれらの位置情報とともに次第に不鮮明になるためである。図1(a)と図1(b)は顔の画像に対して平均化フィルタを適用した結果を示している。同図のごとく、極値をとる画素の輝度が平均化によって次第に薄れるとともに、位置も平均化の影響でシフトしていく。その結果、目（輝度の極小点）の輝度や位置の情報は、このような粗い解像度レベルで曖昧になり、この解像度では正しいマッチングを計算することができない。したがって、粗い解像度レベルを設けるのが大域的なマッチングのためでありながら、ここで得られたマッチングは画像の本当の特徴（目、つま

り極小点)に正確に対応しない。より精細な解像度レベルで目が鮮明に現れたとしても、大域的なマッチングをとる際に混入した誤差は、もはや取り返しがつかない。入力画像にスムージング処理を加えることにより、テクスチャ領域のステレオ情報が落ちてしまうこともすでに指摘されている。

【0021】

一方、最近地形学の分野で利用されはじめた非線形フィルタとして一次元の「ふるい (sieve)」演算子がある。この演算子は、所定の大きさの一次元ウィンドウ内の極小値 (または極大値) を選択することにより、縮尺と空間の因果関係を保存しながら画像にスムージング処理を加える。その結果得られる画像は元の画像と同じ大きさであるが、小さな波の成分が取り除かれているため、より単純になる。画像の情報を落とすという点で、この演算子は広い意味での「多重解像度フィルタ」に分類することはできるが、実際にはウェーブレットのように画像の解像度を変えながら画像を階層化するわけではなく (つまり狭い意味での多重解像度フィルタではなく)、画像間の対応の検出には利用できない。

【0022】

[前提技術が解決しようとする課題]

以上をまとめれば以下の課題が認められる。

1. 画像の特徴を正確に、かつ比較的簡単な処理で把握する画像処理方法が乏しかった。特に、特徴のある点に関する情報、例えば画素値や位置を維持しながら特徴を抽出できる画像処理方法に関する有効な提案が少なかった。

2. 画像の特徴をもとに対応点を自動検出する場合、一般に処理が複雑であるか、ノイズ耐性が低いなどの欠点があった。また、処理に際していろいろな制約を設ける必要があり、大域的特徴及び局所的特徴を同時に満たすマッチングをとることが困難だった。

3. 画像の大域的な構造または特徴を認識するために多重解像度フィルタを導入しても、そのフィルタが線形フィルタの場合、画素の輝度情報と位置情報が曖昧になった。その結果、対応点の把握が不正確になりやすかった。非線形フィルタである一次元ふるい演算子は画像を階層化しないため、画像間の対応点の検出には利用できなかった。

4. これらの結果、対応点を正しく把握しようとするれば、結局人手による指定に頼るほか有効な手だてがなかった。

【 0 0 2 3 】

前提技術はこれらの課題の解決を目的としてなされたものであり、画像処理の分野において、画像の特徴の的確な把握を可能にする技術を提供するものである。

【 0 0 2 4 】

[前提技術が課題を解決するための手段]

この目的のために前提技術のある態様は、新たな多重解像度の画像フィルタを提案する。この多重解像度フィルタは画像から特異点を抽出する。したがって、特異点フィルタともよばれる。特異点とは画像上特徴をもつ点をいう。例として、ある領域において画素値（画素値とは、色番号、輝度値など画像または画素に関する任意の数値を指す）が最大になる極大点、最小になる極小点、ある方向については最大だが別の方向については最小になるような鞍点がある。特異点は位相幾何学上の概念であってもよい。ただし、その他どのような特徴を有してもよい。いかなる性質の点を特異点と考えるかは、前提技術にとって本質問題ではない。

【 0 0 2 5 】

この態様では、多重解像度フィルタを用いた画像処理が行われる。まず検出工程において、第一の画像に対し、二次元的な探索を行って特異点が検出される。つぎに生成工程において、検出された特異点を抽出して第一の画像よりも解像度の低い第二の画像が生成される。第二の画像には第一の画像のもつ特異点が引き継がれる。第二の画像は第一の画像よりも解像度が低いため、画像の大域的な特徴の把握に好適である。

【 0 0 2 6 】

前提技術の別の態様は特異点フィルタを用いた画像マッチング方法に関する。この態様では、始点画像と終点画像間のマッチングがとられる。始点画像および終点画像とは、ふたつの画像の区別のために便宜的に与えた名称であり、本質的な違いはない。

【0027】

この態様では、第一工程にて、始点画像に特異点フィルタを施して解像度の異なる一連の始点階層画像が生成される。第二工程では、終点画像に特異点フィルタを施して解像度の異なる一連の終点階層画像が生成される。始点階層画像、終点階層画像とは、それぞれ始点画像、終点画像を階層化して得られる画像群をいい、それぞれ最低2枚の画像からなる。つぎに第三工程において、始点階層画像と終点階層画像のマッチングが解像度レベルの階層の中で計算される。この態様によれば、多重解像度フィルタによって特異点に関連する画像の特徴が抽出され、および／または明確化されるため、マッチングが容易になる。マッチングのための拘束条件は特に必要としない。

前提技術のさらに別の態様も始点画像と終点画像のマッチングに関する。この態様では、予め複数のマッチング評価項目のそれぞれに関して評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索する。総合評価式は、評価式の少なくともひとつに係数パラメータを掛けたうえでそれらの評価式の総和として定義してもよく、その場合、総合評価式またはいずれかの評価式がほぼ極値をとる状態を検出して前記パラメータを決定してもよい。「極値付近」または「ほぼ極値をとる」としたのは、多少誤差を含んでいてもよいためである。多少の誤差は前提技術にはさして問題とならない。

極値自体も前記パラメータに依存するため、極値の挙動、つまり極値の変化の様子をもとに、最適と考えられるパラメータを決定する余地が生じる。この態様はその事実を利用している。この態様によれば、元来調整の困難なパラメータの決定を自動化する途が拓かれる。

【0028】

[前提技術の実施の形態]

最初に〔1〕で前提技術の要素技術の詳述し、〔2〕で処理手順を具体的に説明する。

〔1〕要素技術の詳細

〔1.1〕イントロダクション

特異点フィルタと呼ばれる新たな多重解像度フィルタを導入し、画像間のマッチングを正確に計算する。オブジェクトに関する予備知識は一切不要である。画像間のマッチングの計算は、解像度の階層を進む間、各解像度において計算される。その際、粗いレベルから精細なレベルへと順に解像度の階層を辿っていく。計算に必要なパラメータは、人間の視覚システムに似た動的計算によって完全に自動設定される。画像間の対応点を人手で特定する必要はない。

【 0 0 2 9 】

本前提技術は、例えば完全に自動的なモーフィング、物体認識、立体写真測量、ボリウムレンダリング、少ないフレームからの滑らかな動画像の生成などに応用できる。モーフィングに用いる場合、与えられた画像を自動的に変形することができる。ボリウムレンダリングに用いる場合、断面間の中間的な画像を正確に再構築することができる。断面間の距離が遠く、断面の形状が大きく変化する場合でも同様である。

【 0 0 3 0 】

〔 1 . 2 〕 特異点フィルタの階層

前提技術に係る多重解像度特異点フィルタは、画像の解像度を落としながら、しかも画像に含まれる各特異点の輝度及び位置を保存することができる。ここで画像の幅を N 、高さを M とする。以下簡単のため、 $N = M = 2^n$ (n は自然数)と仮定する。また、区間 $[0, N] \subset \mathbb{R}$ を I と記述する。 (i, j) における画像の画素を $p(i, j)$ と記述する ($i, j \in I$)。

【 0 0 3 1 】

ここで多重解像度の階層を導入する。階層化された画像群は多重解像度フィルタで生成される。多重解像度フィルタは、もとの画像に対して二次元的な探索を行って特異点を検出し、検出された特異点を抽出してもとの画像よりも解像度の低い別の画像を生成する。ここで第 m レベルにおける各画像のサイズは $2^m \times 2^m$ ($0 \leq m \leq n$) とする。特異点フィルタは次の4種類の新たな階層画像を n から下がる方向で再帰的に構築する。

【数 1】

$$\begin{aligned}
 p_{(i,j)}^{(m,0)} &= \min(\min(p_{(2i,2j)}^{(m+1,0)}, p_{(2i,2j+1)}^{(m+1,0)}), \min(p_{(2i+1,2j)}^{(m+1,0)}, p_{(2i+1,2j+1)}^{(m+1,0)})) \\
 p_{(i,j)}^{(m,1)} &= \max(\min(p_{(2i,2j)}^{(m+1,1)}, p_{(2i,2j+1)}^{(m+1,1)}), \min(p_{(2i+1,2j)}^{(m+1,1)}, p_{(2i+1,2j+1)}^{(m+1,1)})) \\
 p_{(i,j)}^{(m,2)} &= \min(\max(p_{(2i,2j)}^{(m+1,2)}, p_{(2i,2j+1)}^{(m+1,2)}), \max(p_{(2i+1,2j)}^{(m+1,2)}, p_{(2i+1,2j+1)}^{(m+1,2)})) \\
 p_{(i,j)}^{(m,3)} &= \max(\max(p_{(2i,2j)}^{(m+1,3)}, p_{(2i,2j+1)}^{(m+1,3)}), \max(p_{(2i+1,2j)}^{(m+1,3)}, p_{(2i+1,2j+1)}^{(m+1,3)}))
 \end{aligned}$$

(式 1)

ただしここで、

【数 2】

$$p_{(i,j)}^{(n,0)} = p_{(i,j)}^{(n,1)} = p_{(i,j)}^{(n,2)} = p_{(i,j)}^{(n,3)} = p_{(i,j)} \quad (\text{式 2})$$

とする。以降これら 4 つの画像を副画像 (サブイメージ) と呼ぶ。 $\min_{x \leq t} \leq x+1$ 、 $\max_{x \leq t \leq x+1}$ をそれぞれ α 及び β と記述すると、副画像はそれぞれ以下のように記述できる。

【0 0 3 2】

$$\begin{aligned}
 P(m, 0) &= \alpha(x) \alpha(y) p(m+1, 0) \\
 P(m, 1) &= \alpha(x) \beta(y) p(m+1, 1) \\
 P(m, 2) &= \beta(x) \alpha(y) p(m+1, 2) \\
 P(m, 3) &= \beta(x) \beta(y) p(m+1, 3)
 \end{aligned}$$

すなわち、これらは α と β のテンソル積のようなものと考えられる。副画像はそれぞれ特異点に対応している。これらの式から明らかなように、特異点フィルタはもとの画像について 2×2 画素で構成されるブロックごとに特異点を検出する。その際、各ブロックのふたつの方向、つまり縦と横について、最大画素値または最小画素値をもつ点を探索する。画素値として、前提技術では輝度を採用するが、画像に関するいろいろな数値を採用することができる。ふたつの方向の両方について最大画素値となる画素は極大点、ふたつの方向の両方について最小画素値となる画素は極小点、ふたつの方向の一方について最大画素値となるとともに、他方について最小画素値となる画素は鞍点として検出される。

【0033】

特異点フィルタは、各ブロックの内部で検出された特異点の画像（ここでは1画素）でそのブロックの画像（ここでは4画素）を代表させることにより、画像の解像度を落とす。特異点の理論的な観点からすれば、 $\alpha(x)$ $\alpha(y)$ は極小点を保存し、 $\beta(x)$ $\beta(y)$ は極大点を保存し、 $\alpha(x)$ $\beta(y)$ 及び $\beta(x)$ $\alpha(y)$ は鞍点を保存する。

【0034】

はじめに、マッチングをとるべき始点（ソース）画像と終点（デスティネーション）画像に対して別々に特異点フィルタ処理を施し、それぞれ一連の画像群、すなわち始点階層画像と終点階層画像を生成しておく。始点階層画像と終点階層画像は、特異点の種類に対応してそれぞれ4種類ずつ生成される。

【0035】

この後、一連の解像度レベルの中で始点階層画像と終点階層画像のマッチングがとれられていく。まず $p(m, 0)$ を用いて極小点のマッチングがとられる。次に、その結果に基づき、 $p(m, 1)$ を用いて鞍点のマッチングがとられ、 $p(m, 2)$ を用いて他の鞍点のマッチングがとられる。そして最後に $p(m, 3)$ を用いて極大点のマッチングがとられる。

【0036】

図1(c)と図1(d)はそれぞれ図1(a)と図1(b)の副画像 $p(5, 0)$ を示している。同様に、図1(e)と図1(f)は $p(5, 1)$ 、図1(g)と図1(h)は $p(5, 2)$ 、図1(i)と図1(j)は $p(5, 3)$ をそれぞれ示している。これらの図からわかるとおり、副画像によれば画像の特徴部分のマッチングが容易になる。まず $p(5, 0)$ によって目が明確になる。目は顔の中で輝度の極小点だからである。 $p(5, 1)$ によれば口が明確になる。口は横方向で輝度が低いためである。 $p(5, 2)$ によれば首の両側の縦線が明確になる。最後に、 $p(5, 3)$ によって耳や頬の最も明るい点が明確になる。これらは輝度の極大点だからである。

【0037】

特異点フィルタによれば画像の特徴が抽出できるため、例えばカメラで撮影さ

れた画像の特徴と、予め記録しておいたいくつかのオブジェクトの特徴を比較することにより、カメラに映った被写体を識別することができる。

【0038】

[1. 3] 画像間の写像の計算

始点画像の位置 (i, j) の画素を $p^{(n)}(i, j)$ と書き、同じく終点画像の位置 (k, l) の画素を $q^{(n)}(k, l)$ で記述する。 $i, j, k, l \in I$ とする。画像間の写像のエネルギー（後述）を定義する。このエネルギーは、始点画像の画素の輝度と終点画像の対応する画素の輝度の差、及び写像の滑らかさによって決まる。最初に最小のエネルギーを持つ $p^{(m, 0)}$ と $q^{(m, 0)}$ 間の写像 $f^{(m, 0)} : p^{(m, 0)} \rightarrow q^{(m, 0)}$ が計算される。 $f^{(m, 0)}$ に基づき、最小エネルギーを持つ $p^{(m, 1)}$ 、 $q^{(m, 1)}$ 間の写像 $f^{(m, 1)}$ が計算される。この手続は、 $p^{(m, 3)}$ と $q^{(m, 3)}$ の間の写像 $f^{(m, 3)}$ の計算が終了するまで続く。各写像 $f^{(m, i)}$ ($i = 0, 1, 2, \dots$) を副写像と呼ぶことにする。 $f^{(m, i)}$ の計算の都合のために、 i の順序は次式のように並べ替えることができる。並べ替えが必要な理由は後述する。

【0039】

【数3】

$$f^{(m, i)} : p^{(m, \sigma(i))} \rightarrow q^{(m, \sigma(i))} \quad (\text{式3})$$

ここで $\sigma(i) \in \{0, 1, 2, 3\}$ である。

【0040】

[1. 3. 1] 全単射

始点画像と終点画像の間のマッチングを写像で表現する場合、その写像は両画像間で全単射条件を満たすべきである。両画像に概念上の優劣はなく、互いの画素が全射かつ単射で接続されるべきだからである。しかしながら通常の場合とは異なり、ここで構築すべき写像は全単射のデジタル版である。前提技術では、画素は格子点によって特定される。

【0041】

始点副画像（始点画像について設けられた副画像）から終点副画像（終点画像

について設けられた副画像)への写像は、 $f^{(m,s)} : I/2^{n-m} \times I/2^{n-m} \rightarrow I/2^{n-m} \times I/2^{n-m}$ ($s=0, 1, \dots$) によって表される。ここで、 $f^{(m,s)}(i, j) = (k, l)$ は、始点画像の $p^{(m,s)}(i, j)$ が終点画像の $q^{(m,s)}(k, l)$ に写像されることを意味する。簡単のために、 $f(i, j) = (k, l)$ が成り立つとき画素 $q(k, l)$ を $q_f(i, j)$ と記述する。

【0042】

前提技術で扱う画素(格子点)のようにデータが離散的な場合、全単射の定義は重要である。ここでは以下のように定義する(i, i', j, j', k, l は全て整数とする)。まず始めに、始点画像の平面においてRによって表記される各正方形領域、

【数4】

$$p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)}$$

(式4)

を考える($i=0, \dots, 2^m-1, j=0, \dots, 2^m-1$)。ここでRの各辺(エッジ)の方向を以下のように定める。

【数5】

$$\overrightarrow{p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)}}, \overrightarrow{p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)}}, \overrightarrow{p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)}} \text{ and } \overrightarrow{p_{(i,j+1)}^{(m,s)} p_{(i,j)}^{(m,s)}}$$

(式5)

この正方形は写像fによって終点画像平面における四辺形に写像されなければならない。 $f^{(m,s)}(R)$ によって示される四辺形、

【数6】

$$q_{f(i,j)}^{(m,s)} q_{f(i+1,j)}^{(m,s)} q_{f(i+1,j+1)}^{(m,s)} q_{f(i,j+1)}^{(m,s)}$$

(式6)

は、以下の全単射条件を満たす必要がある。

【0043】

1. 四辺形 $f^{(m,s)}(R)$ のエッジは互いに交差しない。
2. $f^{(m,s)}(R)$ のエッジの方向はRのそれらに等しい(図2の場合、時計回り)。

3. 緩和条件として収縮写像（リトラクション：retractions）を許す。

【0044】

何らかの緩和条件を設けないかぎり、全単射条件を完全に満たす写像は単位写像しかないためである。ここでは $f^{(m, s)}(R)$ のひとつのエッジの長さが0、すなわち $f^{(m, s)}(R)$ は三角形になってもよい。しかし、面積が0となるような図形、すなわち1点または1本の線分になってはならない。図2(R)がもとの四辺形の場合、図2(A)と図2(D)は全単射条件を満たすが、図2(B)、図2(C)、図2(E)は満たさない。

【0045】

実際のインプリメンテーションでは、写像が全射であることを容易に保証すべく、さらに以下の条件を課してもよい。つまり始点画像の境界上の各画素は、終点画像において同じ位置を占める画素に写影されるというものである。すなわち、 $f(i, j) = (i, j)$ （ただし $i = 0, i = 2^m - 1, j = 0, j = 2^m - 1$ の4本の線上）である。この条件を以下「付加条件」とも呼ぶ。

【0046】

[1. 3. 2] 写像のエネルギー

[1. 3. 2. 1] 画素の輝度に関するコスト

写像 f のエネルギーを定義する。エネルギーが最小になる写像を探すことが目的である。エネルギーは主に、始点画像の画素の輝度とそれに対応する終点画像の画素の輝度の差で決まる。すなわち、写像 $f^{(m, s)}$ の点 (i, j) におけるエネルギー $C^{(m, s)}(i, j)$ は次式によって定まる。

【数7】

$$C_{(i,j)}^{(m,s)} = |V(p_{(i,j)}^{(m,s)}) - V(q_{f(i,j)}^{(m,s)})|^2 \quad (\text{式7})$$

ここで、 $V(p_{(i,j)}^{(m,s)})$ 及び $V(q_{f(i,j)}^{(m,s)})$ はそれぞれ画素 $p_{(i,j)}^{(m,s)}$ 及び $q_{f(i,j)}^{(m,s)}$ の輝度である。 f のトータルのエネルギー $C^{(m, s)}$ は、マッチングを評価するひとつの評価式であり、つぎに示す $C^{(m, s)}(i, j)$ の合計で定義できる。

【数 8】

$$C_f^{(m,s)} = \sum_{i=0}^{2^m-1} \sum_{j=0}^{2^m-1} C_{(i,j)}^{(m,s)} \quad (\text{式 8})$$

[1. 3. 2. 2] 滑らかな写像のための画素の位置に関するコスト

滑らかな写像を得るために、写像に関する別のエネルギー D_f を導入する。このエネルギーは画素の輝度とは関係なく、 $p^{(m,s)}(i,j)$ および $q^{(m,s)}(i,j)$ の位置によって決まる ($i=0, \dots, 2^m-1, j=0, \dots, 2^m-1$)。点 (i,j) における写像 $f^{(m,s)}$ のエネルギー $D^{(m,s)}(i,j)$ は次式で定義される。

【数 9】

$$D_{(i,j)}^{(m,s)} = \eta E_{0(i,j)}^{(m,s)} + E_{1(i,j)}^{(m,s)} \quad (\text{式 9})$$

ただし、係数パラメータ η は 0 以上の実数であり、また、

【数 10】

$$E_{0(i,j)}^{(m,s)} = \|(i,j) - f^{(m,s)}(i,j)\|^2 \quad (\text{式 10})$$

【数 11】

$$E_{1(i,j)}^{(m,s)} = \sum_{i'=i-1}^i \sum_{j'=j-1}^j \|(f^{(m,s)}(i,j) - (i,j)) - (f^{(m,s)}(i',j') - (i',j'))\|^2 / 4 \quad (\text{式 11})$$

とする。ここで、

【数 12】

$$\|(x,y)\| = \sqrt{x^2 + y^2} \quad (\text{式 12})$$

であり、 $i' < 0$ および $j' < 0$ に対して $f(i', j')$ は 0 と決める。 E_0 は (i,j) 及び $f(i,j)$ の距離で決まる。 E_0 は画素があまりにも離れた画素へ写影されることを防ぐ。ただし E_0 は、後に別のエネルギー関数で置き換える。 E_1 は写像の滑らかさを保証する。 E_1 は、 $p(i,j)$ の変位とその隣接点の変位の間の隔たりを表す。以上の考察をもとに、マッチングを評価する別の評価式であるエネルギー D_f は次式で定まる。

【数 1 3】

$$D_f^{(m,s)} = \sum_{i=0}^{2^m-1} \sum_{j=0}^{2^m-1} D_{(i,j)}^{(m,s)} \quad (\text{式 1 3})$$

[1 . 3 . 2 . 3] 写像の総エネルギー

写像の総エネルギー、すなわち複数の評価式の統合に係る総合評価式は $\lambda C_f^{(m,s)} + D_f^{(m,s)}$ で定義される。ここで係数パラメータ λ は 0 以上の実数である。目的は総合評価式が極値をとる状態を検出すること、すなわち次式で示す最小エネルギーを与える写像を見いだすことである。

【数 1 4】

$$\min_f \lambda C_f^{(m,s)} + D_f^{(m,s)} \quad (\text{式 1 4})$$

$\lambda = 0$ 及び $\eta = 0$ の場合、写像は単位写像になることに注意すべきである（すなわち、全ての $i = 0, \dots, 2^m - 1$ 及び $j = 0, \dots, 2^m - 1$ に対して $f^{(m,s)}(i, j) = (i, j)$ となる）。後述のごとく、本前提技術では最初に $\lambda = 0$ 及び $\eta = 0$ の場合を評価するため、写像を単位写像から徐々に変形していくことができる。仮に総合評価式の λ の位置を変えて $C_f^{(m,s)} + \lambda D_f^{(m,s)}$ と定義したとすれば、 $\lambda = 0$ 及び $\eta = 0$ の場合に総合評価式が $C_f^{(m,s)}$ だけになり、本来何等関連のない画素どうしが単に輝度が近いというだけで対応づけられ、写像が無意味なものになる。そうした無意味な写像をもとに写像を変形していてもまったく意味をなさない。このため、単位写像が評価の開始時点で最良の写像として選択されるよう係数パラメータの与えかたが配慮されている。

【 0 0 4 7 】

オプティカルフローもこの前提技術同様、画素の輝度の差と滑らかさを考慮する。しかし、オプティカルフローは画像の変換に用いることはできない。オブジェクトの局所的な動きしか考慮しないためである。前提技術に係る特異点フィルタを用いることによって大域的な対応関係を検出することができる。

【 0 0 4 8 】

[1 . 3 . 3] 多重解像度の導入による写像の決定

最小エネルギーを与え、全単射条件を満足する写像 f_{min} を多重解像度の階層を用いて求める。各解像度レベルにおいて始点副画像及び終点副画像間の写像を計算する。解像度の階層の最上位（最も粗いレベル）からスタートし、各解像度レベルの写像を、他のレベルの写像を考慮に入れながら決定する。各レベルにおける写像の候補の数は、より高い、つまりより粗いレベルの写像を用いることによって制限される。より具体的には、あるレベルにおける写像の決定に際し、それよりひとつ粗いレベルにおいて求められた写像が一種の拘束条件として課される。

【0049】

まず、

【数15】

$$(i', j') = ([\frac{i}{2}], [\frac{j}{2}]) \quad (\text{式15})$$

が成り立つとき、 $p^{(m-1, s)}(i', j')$ 、 $q^{(m-1, s)}(i', j')$ をそれぞれ $p^{(m, s)}(i, j)$ 、 $q^{(m, s)}(i, j)$ の parent と呼ぶことにする。 $[x]$ は x を越えない最大整数である。また $p^{(m, s)}(i, j)$ 、 $q^{(m, s)}(i, j)$ をそれぞれ $p^{(m-1, s)}(i', j')$ 、 $q^{(m-1, s)}(i', j')$ の child と呼ぶ。関数 $parent(i, j)$ は次式で定義される。

【数16】

$$parent(i, j) = ([\frac{i}{2}], [\frac{j}{2}]) \quad (\text{式16})$$

$p^{(m, s)}(i, j)$ と $q^{(m, s)}(k, l)$ の間の写像 $f^{(m, s)}(i, j) = (k, l)$ の値は $f^{(m-1, s)}(i', j')$ ($m=1, 2, \dots, n$) を用いることによって、以下のように決定される。まず、 $q^{(m, s)}(k, l)$ は次の四辺形の内部になければならないという条件を課し、全単射条件を満たす写像のうち現実性の高いものを絞り込む。

【数 17】

$$q_{g^{(m,s)}(i-1,j-1)}^{(m,s)} q_{g^{(m,s)}(i-1,j+1)}^{(m,s)} q_{g^{(m,s)}(i+1,j+1)}^{(m,s)} q_{g^{(m,s)}(i+1,j-1)}^{(m,s)}$$

(式 17)

ただしここで、

【数 18】

$$g^{(m,s)}(i,j) = f^{(m-1,s)}(\text{parent}(i,j)) + f^{(m-1,s)}(\text{parent}((i,j) + (1,1)))$$

(式 18)

である。こうして定めた四辺形を、以下 $p^{(m,s)}(i,j)$ の相続 (inherited) 四辺形と呼ぶことにする。相続四辺形の内部において、エネルギーを最小にする画素を求める。

【0050】

図3は以上の手順を示している。同図において、始点画像のA, B, C, Dの画素は、第 $m-1$ レベルにおいてそれぞれ終点画像のA', B', C', D'へ写影される。画素 $p^{(m,s)}(i,j)$ は、相続四辺形A' B' C' D'の内部に存在する画素 $q^{(m,s)}(i,j)$ へ写影されなければならない。以上の配慮により、第 $m-1$ レベルの写像から第 m レベルの写像への橋渡しがなされる。

【0051】

先に定義したエネルギー E_0 は、第 m レベルにおける副写像 $f^{(m,0)}$ を計算するために、次式に置き換える。

【数 19】

$$E_{0(i,j)} = \|f^{(m,0)}(i,j) - g^{(m)}(i,j)\|^2 \quad (\text{式 19})$$

また、副写像 $f^{(m,s)}$ を計算するためには次式を用いる。

【数 20】

$$E_{0(i,j)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2 \quad (1 \leq i) \quad (\text{式 20})$$

こうしてすべての副写像のエネルギーを低い値に保つ写像が得られる。式20により、異なる特異点に対応する副写像が、副写像どうしの類似度が高くなるよ

うに同一レベル内で関連づけられる。式19は、 $f^{(m, s)}(i, j)$ と、第 $m-1$ レベルの画素の一部と考えた場合の (i, j) が射影されるべき点の位置との距離を示している。

【0052】

仮に、相続四辺形 $A' B' C' D'$ の内部に全単射条件を満たす画素が存在しない場合は以下の措置をとる。まず、 $A' B' C' D'$ の境界線からの距離が L (始めは $L=1$) である画素を調べる。それらのうち、エネルギーが最小になるものが全単射条件を満たせば、これを $f^{(m, s)}(i, j)$ の値として選択する。そのような点が発見されるか、または L がその上限の $L^{(m)}_{max}$ に到達するまで、 L を大きくしていく。 $L^{(m)}_{max}$ は各レベル m に対して固定である。そのような点が全く発見されない場合、全単射の第3の条件を一時的に無視して変換先の四辺形の面積がゼロになるような写像も認め、 $f^{(m, s)}(i, j)$ を決定する。それでも条件を満たす点が見つからない場合、つぎに全単射の第1及び第2条件を外す。

【0053】

多重解像度を用いる近似法は、写像が画像の細部に影響されることを回避しつつ、画像間の大域的な対応関係を決定するために必須である。多重解像度による近似法を用いなければ、距離の遠い画素間の対応関係を見いだすことは不可能である。その場合、画像のサイズはきわめて小さなものに限定しなければならず、変化の小さな画像しか扱うことができない。さらに、通常写像に滑らかさを要求するため、そうした画素間の対応関係を見つけにくくしている。距離のある画素から画素への写像のエネルギーは高いためである。多重解像度を用いた近似法によれば、そうした画素間の適切な対応関係を見いだすことができる。それらの距離は、解像度の階層の上位レベル(粗いレベル)において小さいためである。

【0054】

〔1.4〕最適なパラメータ値の自動決定

既存のマッチング技術の主な欠点のひとつに、パラメータ調整の困難さがある。大抵の場合、パラメータの調整は人手作業によって行われ、最適な値を選択することはきわめて難しい。前提技術に係る方法によれば、最適なパラメータ値を

完全に自動決定することができる。

【0055】

前提技術に係るシステムはふたつのパラメータ、 λ 及び η を含む。端的にいえば、 λ は画素の輝度の差の重みであり、 η は写像の剛性を示している。これらのパラメータの値は初期値が 0 であり、まず $\eta = 0$ に固定して λ を 0 から徐々に増加させる。 λ の値を大きくしながら、しかも総合評価式（式 14）の値を最小にする場合、各副写像に関する $C(m, s)_f$ の値は一般に小さくなっていく。このことは基本的にふたつの画像がよりマッチしなければならないことを意味する。しかし、 λ が最適値を超えると以下の現象が発生する。

【0056】

1. 本来対応すべきではない画素どうしが、単に輝度が近いというだけで誤って対応づけられる。

2. その結果、画素どうしの対応関係がおかしくなり、写像がくずれはじめる。

3. その結果、式 14 において $D(m, s)_f$ が急激に増加しようとする。

4. その結果、式 14 の値が急激に増加しようとするため、 $D(m, s)_f$ の急激な増加を抑制するよう $f(m, s)$ が変化し、その結果 $C(m, s)_f$ が増加する。

したがって、 λ を増加させながら式 14 が最小値をとるという状態を維持しつつ $C(m, s)_f$ が減少から増加に転じる閾値を検出し、その λ を $\eta = 0$ における最適値とする。つぎに η を少しずつ増やして $C(m, s)_f$ の挙動を検査し、後述の方法で η を自動決定する。その η に対応して λ も決まる。

【0057】

この方法は、人間の視覚システムの焦点機構の動作に似ている。人間の視覚システムでは、一方の目を動かしながら左右両目の画像のマッチングがとられる。オブジェクトがはっきりと認識できるとき、その目が固定される。

【0058】

[1. 4. 1] λ の動的決定

λ は 0 から所定の刻み幅で増加されていき、 λ の値が変わる度に副写像が評価

される。式14のごとく、総エネルギーは $\lambda C^{(m,s)}_f + D^{(m,s)}_f$ に
よって定義される。式9の $D^{(m,s)}_f$ は滑らかさを表すもので、理論的には
単位写像の場合に最小になり、写像が歪むほど E_0 も E_1 も増加していく。 E_1
は整数であるから、 $D^{(m,s)}_f$ の最小刻み幅は1である。このため、現在の
 $\lambda C^{(m,s)}_{(i,j)}$ の変化(減少量)が1以上でなければ、写像を変化さ
せることによって総エネルギーを減らすことはできない。なぜなら、写像の変化
に伴って $D^{(m,s)}_f$ は1以上増加するため、 $\lambda C^{(m,s)}_{(i,j)}$ が1
以上減少しない限り総エネルギーは減らないためである。

【0059】

この条件のもと、 λ の増加に伴い、正常な場合に $C^{(m,s)}_{(i,j)}$ が減
少することを示す。 $C^{(m,s)}_{(i,j)}$ のヒストグラムを $h(l)$ と記述す
る。 $h(l)$ はエネルギー $C^{(m,s)}_{(i,j)}$ が l^2 である画素の数である
。 $\lambda l^2 \geq 1$ が成り立つために、例えば $l^2 = 1/\lambda$ の場合を考える。 λ が λ_1
から λ_2 まで微小量変化するとき、

【数21】

$$A = \sum_{l=\lceil \frac{1}{\lambda_2} \rceil}^{\lfloor \frac{1}{\lambda_1} \rfloor} h(l) \simeq \int_{\frac{1}{\lambda_2}}^{\frac{1}{\lambda_1}} h(l) dl = - \int_{\lambda_2}^{\lambda_1} h(l) \frac{1}{\lambda^{3/2}} d\lambda = \int_{\lambda_1}^{\lambda_2} \frac{h(l)}{\lambda^{3/2}} d\lambda$$

(式21)

で示されるA個の画素が、

【数22】

$$C_f^{(m,s)} - l^2 = C_f^{(m,s)} - \frac{1}{\lambda} \quad (\text{式22})$$

のエネルギーを持つより安定的な状態に変化する。ここでは仮に、これらの画素
のエネルギーがすべてゼロになると近似している。この式は $C^{(m,s)}_f$ の値
が、

【数23】

$$\partial C_f^{(m,s)} = -\frac{A}{\lambda} \quad (\text{式23})$$

だけ変化することを示し、その結果、

【数24】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{h(l)}{\lambda^{5/2}} \quad (\text{式24})$$

が成立する。 $h(1) > 0$ であるから、通常 $C^{(m,s)}_f$ は減少する。しかし、 λ が最適値を越えようとするとき、上述の現象、つまり $C^{(m,s)}_f$ の増加が発生する。この現象を検出することにより、 λ の最適値を決定する。

【0060】

なお、 $H (h > 0)$ 及び k を定数とすると、

【数25】

$$h(l) = H l^k = \frac{H}{\lambda^{k/2}} \quad (\text{式25})$$

と仮定すれば、

【数26】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{H}{\lambda^{5/2+k/2}} \quad (\text{式26})$$

が成り立つ。このとき $k \neq -3$ であれば、

【数27】

$$C_f^{(m,s)} = C + \frac{H}{(3/2 + k/2)\lambda^{3/2+k/2}} \quad (\text{式27})$$

となる。これが $C^{(m,s)}_f$ の一般式である(C は定数)。

【0061】

λ の最適値を検出する際、さらに安全を見て、全単射条件を破る画素の数を検査してもよい。ここで各画素の写像を決定する際、全単射条件を破る確率を p_0 と仮定する。この場合、

【数28】

$$\frac{\partial A}{\partial \lambda} = \frac{h(l)}{\lambda^{3/2}} \quad (\text{式28})$$

が成立しているため、全単射条件を破る画素の数は次式の率で増加する。

【数 29】

$$B_0 = \frac{h(l)p_0}{\lambda^{3/2}}$$

(式 29)

従って、

【数 30】

$$\frac{B_0 \lambda^{3/2}}{p_0 h(l)} = 1$$

(式 30)

は定数である。仮に $h(l) = H l^k$ を仮定するとき、例えば、

【数 31】

$$B_0 \lambda^{3/2+k/2} = p_0 H$$

(式 31)

は定数になる。しかし λ が最適値を越えると、上の値は急速に増加する。この現象を検出し、 $B_0 \lambda^{3/2+k/2} / 2^m$ の値が異常値 $B_{0 \text{ thres}}$ を越えるかどうかを検査し、 λ の最適値を決定することができる。同様に、 $B_1 \lambda^{3/2+k/2} / 2^m$ の値が異常値 $B_{1 \text{ thres}}$ を越えるかどうかを検査することにより、全単射の第3の条件を破る画素の増加率 B_1 を確認する。ファクター 2^m を導入する理由は後述する。このシステムはこれら2つの閾値に敏感ではない。これらの閾値は、エネルギー $C(m, s)_f$ の観察では検出し損なった写像の過度の歪みを検出するために用いることができる。

【0062】

なお実験では、副写像 $f(m, s)$ を計算する際、もし λ が 0.1 を越えたら $f(m, s)$ の計算は止めて $f(m, s+1)$ の計算に移行した。 $\lambda > 0.1$ のとき、画素の輝度 255 レベル中のわずか「3」の違いが副写像の計算に影響したためであり、 $\lambda > 0.1$ のとき正しい結果を得ることは困難だったためである。

【0063】

[1. 4. 2] ヒストグラム $h(l)$

$C(m, s)_f$ の検査はヒストグラム $h(l)$ に依存しない。全単射及びその第3の条件の検査の際、 $h(l)$ に影響を受けうる。実際に $(\lambda, C(m, s)_f)$ をプロットすると、 k は通常 1 付近にある。実験では $k = 1$ を用い、 $B_0 \lambda$

2 と $B_1 \lambda^2$ を検査した。仮に k の本当の値が 1 未満であれば、 $B_0 \lambda^2$ と $B_1 \lambda^2$ は定数にならず、ファクター $\lambda^{(1-k)/2}$ に従って徐々に増加する。 $h(1)$ が定数であれば、例えばファクターは $\lambda^{1/2}$ である。しかし、こうした差は閾値 B_{0thres} を正しく設定することによって吸収することができる。

【0064】

ここで次式のごとく始点画像を中心が (x_0, y_0) 、半径 r の円形のオブジェクトであると仮定する。

【数 3 2】

$$p(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_0)^2 + (j-y_0)^2}) & (\sqrt{(i-x_0)^2 + (j-y_0)^2} \leq r) \\ 0 & (otherwise) \end{cases} \quad (式 3 2)$$

一方、終点画像は、次式のごとく中心 (x_1, y_1) 、半径が r のオブジェクトであるとする。

【数 3 3】

$$q(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_1)^2 + (j-y_1)^2}) & (\sqrt{(i-x_1)^2 + (j-y_1)^2} \leq r) \\ 0 & (otherwise) \end{cases} \quad (式 3 3)$$

ここで $c(x)$ は $c(x) = x^k$ の形であるとする。中心 (x_0, y_0) 及び (x_1, y_1) が十分遠い場合、ヒストグラム $h(1)$ は次式の形となる。

【数 3 4】

$$h(l) \propto rl^k \quad (k \neq 0) \quad (式 3 4)$$

$k = 1$ のとき、画像は背景に埋め込まれた鮮明な境界線を持つオブジェクトを示す。このオブジェクトは中心が暗く、周囲にいくに従って明るくなる。 $k = -1$ のとき、画像は曖昧な境界線を持つオブジェクトを表す。このオブジェクトは中心が最も明るく、周囲にいくに従って暗くなる。一般のオブジェクトはこれらふたつのタイプのオブジェクトの間にあると考えてもさして一般性を失わない。したがって、 k は $-1 \leq k \leq 1$ として大抵の場合をカバーでき、式 27 が一般に減少関数であることが保障される。

【0065】

なお、式 34 からわかるように、 r は画像の解像度に影響されること、すなわち r は 2^m に比例することに注意すべきである。このために [1. 4. 1] に

いてファクター 2^m を導入した。

【0066】

[1. 4. 3] η の動的決定

パラメータ η も同様の方法で自動決定できる。はじめに $\eta = 0$ とし、最も細かい解像度における最終的な写像 $f^{(n)}$ 及びエネルギー $C_f^{(n)}$ を計算する。つづいて、 η をある値 $\Delta \eta$ だけ増加させ、再び最も細かい解像度における最終写像 $f^{(n)}$ 及びエネルギー $C_f^{(n)}$ を計算し直す。この過程を最適値が求まるまで続ける。 η は写像の剛性を示す。次式の重みだからである。

【数35】

$$E_{0(i,j)}^{(m,s)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2$$

(式35)

η が 0 のとき、 $D_f^{(n)}$ は直前の副写像と無関係に決定され、現在の副写像は弾性的に変形され、過度に歪むことになる。一方、 η が非常に大きな値のとき、 $D_f^{(n)}$ は直前の副写像によってほぼ完全に決まる。このとき副写像は非常に剛性が高く、画素は同じ場所に射影される。その結果、写像は単位写像になる。 η の値が 0 から次第に増えるとき、後述のごとく $C_f^{(n)}$ は徐々に減少する。しかし η の値が最適値を越えると、図 4 に示すとおり、エネルギーは増加し始める。同図の X 軸は η 、Y 軸は C_f である。

【0067】

この方法で $C_f^{(n)}$ を最小にする最適な η の値を得ることができる。しかし、 λ の場合に比べていろいろな要素が計算に影響する結果、 $C_f^{(n)}$ は小さく揺らぎながら変化する。 λ の場合は、入力が微小量変化するたびに副写像を 1 回計算しなおすだけだが、 η の場合はすべての副写像が計算しなおされるためである。このため、得られた $C_f^{(n)}$ の値が最小であるかどうかを即座に判断することはできない。最小値の候補が見つければ、さらに細かい区間を設定することによって真の最小値を探す必要がある。

【0068】

[1. 5] スーパーサンプリング

画素間の対応関係を決定する際、自由度を増やすために、 $f(m, s)$ の値域を $R \times R$ に拡張することができる (R は実数の集合)。この場合、終点画像の画素の輝度が補間され、非整数点、

【数 3 6】

$$V(q_{f(m,s)}^{(m,s)}(i,j)) \quad (\text{式 3 6})$$

における輝度を持つ $f(m, s)$ が提供される。つまりスーパーサンプリングが行われる。実験では、 $f(m, s)$ は整数及び半整数値をとることが許され、

【数 3 7】

$$V(q_{(i,j)+(0.5,0.5)}^{(m,s)}) \quad (\text{式 3 7})$$

は、

【数 3 8】

$$(V(q_{(i,j)}^{(m,s)}) + V(q_{(i,j)+(1,1)}^{(m,s)}))/2 \quad (\text{式 3 8})$$

によって与えられた。

【0069】

[1. 6] 各画像の画素の輝度の正規化

始点画像と終点画像がきわめて異なるオブジェクトを含んでいるとき、写像の計算に元の画素の輝度がそのままでは利用しにくい。輝度の差が大きいために輝度に関するエネルギー $C(m, s)_f$ が大きくなりすぎ、正しい評価がしづらいためである。

【0070】

例えば、人の顔と猫の顔のマッチングをとる場合を考える。猫の顔は毛で覆われており、非常に明るい画素と非常に暗い画素が混じっている。この場合、ふたつの顔の間の副写像を計算するために、まず副画像を正規化する。すなわち、最も暗い画素の輝度を 0、最も明るいそれを 255 に設定し、他の画素の輝度は線形補間によって求めておく。

【0071】

[1. 7] インプリメンテーション

始点画像のスキャンに従って計算がリニアに進行する帰納的な方法を用いる。始めに、1番上の左端の画素 $(i, j) = (0, 0)$ について $f(m, s)$ の値を決定する。次に i を1ずつ増やしながらか $f(m, s)(i, j)$ の値を決定する。 i の値が画像の幅に到達したとき、 j の値を1増やし、 i を0に戻す。以降、始点画像のスキャンに伴い $f(m, s)(i, j)$ を決定していく。すべての点について画素の対応が決まれば、ひとつの写像 $f(m, s)$ が決まる。

ある $p(i, j)$ について対応点 $q_f(i, j)$ が決まれば、つぎに $p(i, j+1)$ の対応点 $q_f(i, j+1)$ が決められる。この際、 $q_f(i, j+1)$ の位置は全単射条件を満たすために、 $q_f(i, j)$ の位置によって制限される。したがって、先に対応点が決まる点ほどこのシステムでは優先度が高くなる。つねに $(0, 0)$ が最も優先される状態がずっと、求められる最終の写像に余計な偏向が加わる。本前提技術ではこの状態を回避するために、 $f(m, s)$ を以下の方法で決めていく。

【0072】

まず $(s \bmod 4)$ が0の場合、 $(0, 0)$ を開始点とし i 及び j を徐々に増やしながらか決めていく。 $(s \bmod 4)$ が1の場合、最上行の右端点を開始点とし、 i を減少、 j を増加させながらか決めていく。 $(s \bmod 4)$ が2のとき、最下行の右端点を開始点とし、 i 及び j を減少させながらか決めていく。 $(s \bmod 4)$ が3の場合、最下行の左端点を開始点とし、 i を増加、 j を減少させながらか決めていく。解像度が最も細かい第 n レベルには副写像という概念、すなわちパラメータ s が存在しないため、仮に $s = 0$ 及び $s = 2$ であるとしてふたつの方向を連続的に計算した。

【0073】

実際のインプリメンテーションでは、全単射条件を破る候補に対してペナルティを与えることにより、候補 $(k, 1)$ の中からできる限り全単射条件を満たす $f(m, s)(i, j)$ ($m = 0, \dots, n$) の値を選んだ。第3の条件を破る候補のエネルギー $D(k, 1)$ には ϕ を掛け、一方、第1または第2の条件を破る候補には ψ を掛ける。今回は $\phi = 2$ 、 $\psi = 100000$ を用いた。

【0074】

前述の全単射条件のチェックのために、実際の手続として $(k, 1) = f^{(m, s)}(i, j)$ を決定する際に以下のテストを行った。すなわち $f^{(m, s)}(i, j)$ の相続四辺形に含まれる各格子点 $(k, 1)$ に対し、次式の外積の z 成分が 0 以上になるかどうかを確かめる。

【数 39】

$$W = \vec{A} \times \vec{B} \quad (\text{式 39})$$

ただしここで、

【数 40】

$$\vec{A} = \overrightarrow{q_{f^{(m,s)}(i,j-1)}^{(m,s)} q_{f^{(m,s)}(i+1,j-1)}^{(m,s)}} \quad (\text{式 40})$$

【数 41】

$$\vec{B} = \overrightarrow{q_{f^{(m,s)}(i,j-1)}^{(m,s)} q_{(k,l)}^{(m,s)}}$$

(式 41)

である（ここでベクトルは三次元ベクトルとし、 z 軸は直交右手座標系において定義される）。もし W が負であれば、その候補については $D^{(m, s)}(k, 1)$ に ϕ を掛けることによってペナルティを与え、できるかぎり選択しないようにする。

【0075】

図 5 (a)、図 5 (b) はこの条件を検査する理由を示している。図 5 (a) はペナルティのない候補、図 5 (b) はペナルティがある候補をそれぞれ表す。隣接画素 $(i, j+1)$ に対する写像 $f^{(m, s)}(i, j+1)$ を決定する際、 W の z 成分が負であれば始点画像平面上において全単射条件を満足する画素は存在しない。なぜなら、 $q^{(m, s)}(k, 1)$ は隣接する四辺形の境界線を越えるためである。

【0076】

[1. 7. 1] 副写像の順序

インプリメンテーションでは、解像度レベルが偶数のときには $\sigma(0) = 0$ 、

$\sigma(1) = 1$ 、 $\sigma(2) = 2$ 、 $\sigma(3) = 3$ 、 $\sigma(4) = 0$ を用い、奇数のときは $\sigma(0) = 3$ 、 $\sigma(1) = 2$ 、 $\sigma(2) = 1$ 、 $\sigma(3) = 0$ 、 $\sigma(4) = 3$ を用いた。このことで、副写像を適度にシャッフルした。なお、本来副写像は4種類であり、 s は0～3のいずれかである。しかし、実際には $s = 4$ に相当する処理を行った。その理由は後述する。

【0077】

[1. 8] 補間計算

始点画像と終点画像の間の写像が決定された後、対応しあう画素の輝度が補間される。実験では、トライリニア補間を用いた。始点画像平面における正方形 $p(i, j) \ p(i+1, j) \ p(i, j+1) \ p(i+1, j+1)$ が終点画像平面上の四辺形 $q_f(i, j) \ q_f(i+1, j) \ q_f(i, j+1) \ q_f(i+1, j+1)$ に射影されると仮定する。簡単のため、画像間の距離を1とする。始点画像平面からの距離が t ($0 \leq t \leq 1$) である中間画像の画素 $r(x, y, t)$ ($0 \leq x \leq N-1$, $0 \leq y \leq M-1$) は以下の要領で求められる。まず画素 $r(x, y, t)$ の位置 (ただし $x, y, t \in R$) を次式で求める。

【数42】

$$\begin{aligned} (x, y) = & (1-dx)(1-dy)(1-t)(i, j) + (1-dx)(1-dy)tf(i, j) \\ & + dx(1-dy)(1-t)(i+1, j) + dx(1-dy)tf(i+1, j) \\ & + (1-dx)dy(1-t)(i, j+1) + (1-dx)dytf(i, j+1) \\ & + dx dy(1-t)(i+1, j+1) + dx dytf(i+1, j+1) \end{aligned}$$

(式42)

つづいて $r(x, y, t)$ における画素の輝度が次の式を用いて決定される。

【数43】

$$\begin{aligned} V(r(x, y, t)) = & (1-dx)(1-dy)(1-t)V(p_{(i,j)}) + (1-dx)(1-dy)tV(q_{f(i,j)}) \\ & + dx(1-dy)(1-t)V(p_{(i+1,j)}) + dx(1-dy)tV(q_{f(i+1,j)}) \\ & + (1-dx)dy(1-t)V(p_{(i,j+1)}) + (1-dx)dytV(q_{f(i,j+1)}) \\ & + dx dy(1-t)V(p_{(i+1,j+1)}) + dx dytV(q_{f(i+1,j+1)}) \end{aligned}$$

(式43)

ここで dx 及び dy はパラメータであり、0から1まで変化する。

【0078】

[1. 9] 拘束条件を課したときの写像

いままでは拘束条件がいったい存在しない場合の写像の決定を述べた。しかし、始点画像と終点画像の特定の画素間に予め対応関係が規定されているとき、これを拘束条件としたうえで写像を決定することができる。

【0079】

基本的な考えは、まず始点画像の特定の画素を終点画像の特定の画素に移す大まかな写像によって始点画像を大まかに変形し、しかる後、写像 f を正確に計算する。

【0080】

まず始めに、始点画像の特定の画素を終点画像の特定の画素に射影し、始点画像の他の画素を適当な位置に射影する大まかな写像を決める。すなわち、特定の画素に近い画素は、その特定の画素が射影される場所の近くに射影されるような写像である。ここで第 m レベルの大まかな写像を $F^{(m)}$ と記述する。

【0081】

大まかな写像 F は以下の要領で決める。まず、いくつかの画素について写像を特定する。始点画像について n_s 個の画素、

【数44】

$$P(i_0, j_0), P(i_1, j_1), \dots, P(i_{n_s-1}, j_{n_s-1})$$

(式44)

を特定するとき、以下の値を決める。

【数45】

$$\begin{aligned} F^{(n)}(i_0, j_0) &= (k_0, l_0), \\ F^{(n)}(i_1, j_0) &= (k_1, l_1), \dots, \\ F^{(n)}(i_{n_s-1}, j_{n_s-1}) &= (k_{n_s-1}, l_{n_s-1}) \end{aligned}$$

(式45)

始点画像の他の画素の変位量は、 $P(i_h, j_h)$ ($h = 0, \dots, n_s - 1$) の変位に重み付けをして求められる平均である。すなわち画素 $P(i, j)$ は、

終点画像の以下の画素に射影される。

【数 4 6】

$$F^{(m)}(i, j) = \frac{(i, j) + \sum_{h=0}^{h=n_s-1} (k_h - i_h, l_h - j_h) \text{weight}_h(i, j)}{2^{n-m}} \quad (\text{式 4 6})$$

ただしここで、

【数 4 7】

$$\text{weight}_h(i, j) = \frac{1/\|(i_h - i, j_h - j)\|^2}{\text{total weight}(i, j)} \quad (\text{式 4 7})$$

【数 4 8】

$$\text{total weight}(i, j) = \sum_{h=0}^{h=n_s-1} 1/\|(i_h - i, j_h - j)\|^2 \quad (\text{式 4 8})$$

とする。

【0082】

つづいて、 $F^{(m)}$ に近い候補写像 f がより少ないエネルギーを持つように、その写像 f のエネルギー $D^{(m, s)}(i, j)$ を変更する。正確には、 $D^{(m, s)}(i, j)$ は、

【数 4 9】

$$D_{(i, j)}^{(m, s)} = E_{0(i, j)}^{(m, s)} + \eta E_{1(i, j)}^{(m, s)} + \kappa E_{2(i, j)}^{(m, s)} \quad (\text{式 4 9})$$

である。ただし、

【数 5 0】

$$E_{2(i, j)}^{(m, s)} = \begin{cases} 0, & \text{if } \|F^{(m)}(i, j) - f^{(m, s)}(i, j)\|^2 \leq \lfloor \frac{\rho^2}{2^{2(n-m)}} \rfloor \\ \|F^{(m)}(i, j) - f^{(m, s)}(i, j)\|^2, & \text{otherwise} \end{cases} \quad (\text{式 5 0})$$

であり、 $\kappa, \rho \geq 0$ とする。最後に、前述の写像の自動計算プロセスにより、 f を完全に決定する。

【0083】

ここで、 $f^{(m, s)}(i, j)$ が $F^{(m)}(i, j)$ に十分近いとき、つまりそれらの距離が、

【数 5 1】

$$\left[\frac{\rho^2}{2^{2(n-m)}} \right] \quad (\text{式 5 1})$$

以内であるとき、 $E_2^{(m, s)}(i, j)$ が 0 になることに注意すべきである。そのように定義した理由は、各 $f^{(m, s)}(i, j)$ が $F^{(m)}(i, j)$ に十分近い限り、終点画像において適切な位置に落ち着くよう、その値を自動的に決めたいためである。この理由により、正確な対応関係を詳細に特定する必要がなく、始点画像は終点画像にマッチするように自動的にマッピングされる。

〔2〕具体的な処理手順

〔1〕の各要素技術による処理の流れを説明する。

図 6 は前提技術の全体手順を示すフローチャートである。同図のごとく、まず多重解像度特異点フィルタを用いた処理を行い (S 1)、つづいて始点画像と終点画像のマッチングをとる (S 2)。ただし、S 2 は必須ではなく、S 1 で得られた画像の特徴をもとに画像認識などの処理を行ってもよい。

【0084】

図 7 は図 6 の S 1 の詳細を示すフローチャートである。ここでは S 2 で始点画像と終点画像のマッチングをとることを前提としている。そのため、まず特異点フィルタによって始点画像の階層化を行い (S 10)、一連の始点階層画像を得る。つづいて同様の方法で終点画像の階層化を行い (S 11)、一連の終点階層画像を得る。ただし、S 10 と S 11 の順序は任意であるし、始点階層画像と終点階層画像を並行して生成していくこともできる。

【0085】

図 8 は図 7 の S 10 の詳細を示すフローチャートである。もとの始点画像のサイズは $2^n \times 2^n$ とする。始点階層画像は解像度が細かいほうから順に作られるため、処理の対象となる解像度レベルを示すパラメータ m を n にセットする (S 100)。つづいて第 m レベルの画像 $p^{(m, 0)}$ 、 $p^{(m, 1)}$ 、 $p^{(m, 2)}$ 、 $p^{(m, 3)}$ から特異点フィルタを用いて特異点を検出し (S 101)、そ

れぞれ第 $m-1$ レベルの画像 $p(m-1, 0)$ 、 $p(m-1, 1)$ 、 $p(m-1, 2)$ 、 $p(m-1, 3)$ を生成する(S102)。ここでは $m=n$ であるため、 $p(m, 0) = p(m, 1) = p(m, 2) = p(m, 3) = p(n)$ であり、ひとつの始点画像から4種類の副画像が生成される。

【0086】

図9は第 m レベルの画像の一部と、第 $m-1$ レベルの画像の一部の対応関係を示している。同図の数値は各画素の輝度を示す。同図の $p(m, s)$ は $p(m, 0) \sim p(m, 3)$ の4つの画像を象徴するもので、 $p(m-1, 0)$ を生成する場合には、 $p(m, s)$ は $p(m, 0)$ であると考え。[1. 2]で示した規則により、 $p(m-1, 0)$ は例えば同図で輝度を記入したブロックについて、そこに含まれる4画素のうち「3」、 $p(m-1, 1)$ は「8」、 $p(m-1, 2)$ は「6」、 $p(m-1, 3)$ を「10」をそれぞれ取得し、このブロックをそれぞれ取得したひとつの画素で置き換える。したがって、第 $m-1$ レベルの副画像のサイズは $2^{m-1} \times 2^{m-1}$ になる。

【0087】

つづいて m をデクリメントし(図8のS103)、 m が負になっていないことを確認し(S104)、S101に戻ってつぎに解像度の粗い副画像を生成していく。この繰り返し処理の結果、 $m=0$ 、すなわち第0レベルの副画像が生成された時点でS10が終了する。第0レベルの副画像のサイズは 1×1 である。

【0088】

図10はS10によって生成された始点階層画像を $n=3$ の場合について例示している。最初の始点画像のみが4つの系列に共通であり、以降特異点の種類に応じてそれぞれ独立に副画像が生成されていく。なお、図8の処理は図7のS11にも共通であり、同様の手順を経て終点階層画像も生成される。以上で図6のS1による処理が完了する。

【0089】

前提技術では、図6のS2に進むためにマッチング評価の準備をする。図11はその手順を示している。同図のごとく、まず複数の評価式が設定される(S30)。[1. 3. 2. 1]で導入した画素に関するエネルギー $C(m, s)_f$ と

[1. 3. 2. 2] で導入した写像の滑らかさに関するエネルギー $D(m, s)$ f がそれである。つぎに、これらの評価式を統合して総合評価式を立てる (S 3 1)。[1. 3. 2. 3] で導入した総エネルギー $\lambda C(m, s) f + D(m, s) f$ がそれであり、[1. 3. 2. 2] で導入した η を用いれば、

$$\sum \sum (\lambda C(m, s) (i, j) + \eta E_0(m, s) (i, j) + E_1(m, s) (i, j)) \quad (\text{式 5 2})$$

となる。ただし、総和は i, j についてそれぞれ $0, 1 \dots, 2^m - 1$ で計算する。以上でマッチング評価の準備が整う。

【0090】

図 1 2 は図 6 の S 2 の詳細を示すフローチャートである。[1] で述べたごとく、始点階層画像と終点階層画像のマッチングは互いに同じ解像度レベルの画像どうしでとられる。画像間の大域的なマッチングを良好にとるために、解像度が粗いレベルから順にマッチングを計算する。特異点フィルタを用いて始点階層画像および終点階層画像を生成しているため、特異点の位置や輝度は解像度の粗いレベルでも明確に保存されており、大域的なマッチングの結果は従来に比べて非常に優れたものになる。

【0091】

図 1 2 のごとく、まず係数パラメータ η を 0、レベルパラメータ m を 0 に設定する (S 2 0)。つづいて、始点階層画像中の第 m レベルの 4 つの副画像と終点階層画像中の第 m レベルの 4 つの副画像のそれぞれの間でマッチングを計算し、それぞれ全単射条件を満たし、かつエネルギーを最小にするような 4 種類の副写像 $f(m, s)$ ($s = 0, 1, 2, 3$) を求める (S 2 1)。全単射条件は [1. 3. 3] で述べた相続四辺形を用いて検査される。この際、式 1 7、1 8 が示すように、第 m レベルにおける副写像は第 $m - 1$ レベルのそれらに拘束されるため、より解像度の粗いレベルにおけるマッチングが順次利用されていく。これは異なるレベル間の垂直的参照である。なお、いま $m = 0$ であってそれより粗いレベルはないが、この例外的な処理は図 1 3 で後述する。

一方、同一レベル内における水平的参照も行われる。[1. 3. 3] の式 2 0 のごとく、 $f(m, 3)$ は $f(m, 2)$ に、 $f(m, 2)$ は $f(m, 1)$ に、 f

$(m, 1)$ は $f(m, 0)$ に、それぞれ類似するように決める。その理由は、特異点の種類が違って、それらがもともとと同じ始点画像と終点画像に含まれている以上、副写像がまったく異なるという状況は不自然だからである。式 20 からわかるように、副写像どうしが近いほどエネルギーは小さくなり、マッチングが良好とみなされる。

【0092】

なお、最初に決めるべき $f(m, 0)$ については同一のレベルで参照できる副写像がないため、式 19 に示すごとくひとつ粗いレベルを参照する。ただし、実験では $f(m, 3)$ まで求めた後、これを拘束条件として $f(m, 0)$ を一回更新するという手続をとった。これは式 20 に $s = 4$ を代入し、 $f(m, 4)$ を新たな $f(m, 0)$ とすることに等しい。 $f(m, 0)$ と $f(m, 3)$ の関連度が低くなり過ぎる傾向を回避するためであり、この措置によって実験結果がより良好になった。この措置に加え、実験では [1. 7. 1] に示す副写像のシャッフルも行った。これも本来特異点の種類ごとに決まる副写像どうしの関連度を密接に保つ趣旨である。また、処理の開始点に依存する偏向を回避するために、 s の値にしたがって開始点の位置を変える点は [1. 7] で述べたとおりである。

【0093】

図 13 は第 0 レベルにおいて副写像を決定する様子を示す図である。第 0 レベルでは各副画像がただひとつの画素で構成されるため、4 つの副写像 $f(0, s)$ はすべて自動的に単位写像に決まる。図 14 は第 1 レベルにおいて副写像を決定する様子を示す図である。第 1 レベルでは副画像がそれぞれ 4 画素で構成される。同図ではこれら 4 画素が実線で示されている。いま、 $p(1, s)$ の点 x の対応点を $q(1, s)$ の中で探すとき、以下の手順を踏む。

【0094】

1. 第 1 レベルの解像度で点 x の左上点 a 、右上点 b 、左下点 c 、右下点 d を求める。
2. 点 $a \sim d$ がひとつ粗いレベル、つまり第 0 レベルにおいて属する画素を探す。図 14 の場合、点 $a \sim d$ はそれぞれ画素 $A \sim D$ に属する。ただし、画素 $A \sim C$ は本来存在しない仮想的な画素である。

3. 第0レベルですでに求まっている画素A～Dの対応点A'～D'を $q^{(1, s)}$ の中にプロットする。画素A'～C'は仮想的な画素であり、それぞれ画素A～Cと同じ位置にあるものとする。

4. 画素Aの中の点aの対応点a'が画素A'の中にあるとみなし、点a'をプロットする。このとき、点aが画素Aの中で占める位置（この場合、右下）と、点a'が画素A'の中で占める位置が同じであると仮定する。

5. 4と同様の方法で対応点b'～d'をプロットし、点a'～d'で相続四辺形を作る。

6. 相続四辺形の中でエネルギーが最小になるよう、点xの対応点x'を探す。対応点x'の候補として、例えば画素の中心が相続四辺形に含まれるものに限定してもよい。図14の場合、4つの画素がすべて候補になる。

【0095】

以上がある点xの対応点の決定手順である。同様の処理を他のすべての点について行い、副写像を決める。第2レベル以上のレベルでは、次第に相続四辺形の形が崩れていくと考えられるため、図3に示すように画素A'～D'の間隔が空いていく状況が発生する。

【0096】

こうして、ある第mレベルの4つの副写像が決まれば、mをインクリメントし（図12のS22）、mがnを超えていないことを確かめて（S23）、S21に戻る。以下、S21に戻るたびに次第に細かい解像度のレベルの副写像を求め、最後にS21に戻ったときに第nレベルの写像 $f^{(n)}$ を決める。この写像は $\eta=0$ に関して定まったものであるから、 $f^{(n)} (\eta=0)$ と書く。

【0097】

つぎに異なる η に関する写像も求めるべく、 η を $\Delta\eta$ だけシフトし、mをゼロクリアする（S24）。新たな η が所定の探索打ち切り値 η_{max} を超えていないことを確認し（S25）、S21に戻り、今回の η に関して写像 $f^{(n)} (\eta=\Delta\eta)$ を求める。この処理を繰り返し、S21で $f^{(n)} (\eta=i\Delta\eta)$ ($i=0, 1, \dots$) を求めていく。 η が η_{max} を超えたときS26に進み、後述の方法で最適な $\eta=\eta_{opt}$ を決定し、 $f^{(n)} (\eta=\eta_{opt})$ を最終的に写像f

(n) とする。

【0098】

図15は図12のS21の詳細を示すフローチャートである。このフローチャートにより、ある定まった η について、第 m レベルにおける副写像が決まる。副写像を決める際、前提技術では副写像ごとに最適な λ を独立して決める。

【0099】

同図のごとく、まず s と λ をゼロクリアする(S210)。つぎに、そのときの λ について(および暗に η について)エネルギーを最小にする副写像 $f^{(m, s)}$ を求め(S211)、これを $f^{(m, s)}(\lambda=0)$ と書く。異なる λ に関する写像も求めるべく、 λ を $\Delta\lambda$ だけシフトし、新たな λ が所定の探索打ち切り値 λ_{max} を超えていないことを確認し(S213)、S211に戻り、以降の繰り返し処理で $f^{(m, s)}(\lambda=i\Delta\lambda)$ ($i=0, 1, \dots$)を求める。 λ が λ_{max} を超えたときS214に進み、最適な $\lambda=\lambda_{opt}$ を決定し、 $f^{(m, s)}(\lambda=\lambda_{opt})$ を最終的に写像 $f^{(m, s)}$ とする(S214)。

【0100】

つぎに、同一レベルにおける他の副写像を求めるべく、 λ をゼロクリアし、 s をインクリメントする(S215)。 s が4を超えていないことを確認し(S216)、S211に戻る。 $s=4$ になれば上述のごとく $f^{(m, 3)}$ を利用して $f^{(m, 0)}$ を更新し、そのレベルにおける副写像の決定を終了する。

【0101】

図16は、ある m と s について λ を変えながら求められた $f^{(m, s)}(\lambda=i\Delta\lambda)$ ($i=0, 1, \dots$)に対応するエネルギー $C^{(m, s)}_f$ の挙動を示す図である。[1.4]で述べたとおり、 λ が増加すると通常 $C^{(m, s)}_f$ は減少する。しかし、 λ が最適値を超えると $C^{(m, s)}_f$ は増加に転じる。そこで本前提技術では $C^{(m, s)}_f$ が極小値をとるときの λ を λ_{opt} と決める。同図のように $\lambda>\lambda_{opt}$ の範囲で再度 $C^{(m, s)}_f$ が小さくなっていても、その時点ではすでに写像がくずれていて意味をなさないため、最初の極小点に注目すればよい。 λ_{opt} は副写像ごとに独立して決めていき、最後に $f^{(n)}$ についてもひとつ定まる。

【0102】

一方、図17は、 η を変えながら求められた $f^{(n)}$ ($\eta = i \Delta \eta$) ($i = 0, 1, \dots$) に対応するエネルギー $C^{(n)}_f$ の挙動を示す図である。ここでも η が増加すると通常 $C^{(n)}_f$ は減少するが、 η が最適値を超えると $C^{(n)}_f$ は増加に転じる。そこで $C^{(n)}_f$ が極小値をとるときの η を η_{opt} と決める。図17は図4の横軸のゼロ付近を拡大した図と考えてよい。 η_{opt} が決まれば $f^{(n)}$ を最終決定することができる。

【0103】

以上、本前提技術によれば種々のメリットが得られる。まずエッジを検出する必要がないため、エッジ検出タイプの従来技術の課題を解消できる。また、画像に含まれるオブジェクトに対する先験的な知識も不要であり、対応点の自動検出が実現する。特異点フィルタによれば、解像度の粗いレベルでも特異点の輝度や位置を維持することができ、オブジェクト認識、特徴抽出、画像マッチングに極めて有利である。その結果、人手作業を大幅に軽減する画像処理システムの構築が可能となる。

【0104】

なお、本前提技術について次のような変形技術も考えられる。

(1) 前提技術では始点階層画像と終点階層画像の間でマッチングをとる際にパラメータの自動決定を行ったが、この方法は階層画像間ではなく、通常の2枚の画像間のマッチングをとる場合全般に利用できる。

【0105】

たとえば2枚の画像間で、画素の輝度の差に関するエネルギー E_0 と画素の位置的なずれに関するエネルギー E_1 のふたつを評価式とし、これらの線形和 $E_{tot} = \alpha E_0 + E_1$ を総合評価式とする。この総合評価式の極値付近に注目して α を自動決定する。つまり、いろいろな α について E_{tot} が最小になるような写像を求める。それらの写像のうち、 α に関して E_1 が極小値をとるときの α を最適パラメータと決める。そのパラメータに対応する写像を最終的に両画像間の最適マッチングとみなす。

【0106】

これ以外にも評価式の設定にはいろいろな方法があり、例えば $1/E_1$ と $1/E_2$ のように、評価結果が良好なほど大きな値をとるものを採用してもよい。総合評価式も必ずしも線形和である必要はなく、 n 乗和 ($n=2$ 、 $1/2$ 、 -1 、 -2 など)、多項式、任意の関数などを適宜選択すればよい。

【0107】

パラメータも、 α のみ、前提技術のごとく η と λ のふたつの場合、それ以上の場合など、いずれでもよい。パラメータが3以上の場合はひとつずつ変化させて決めていく。

(2) 本前提技術では、総合評価式の値が最小になるよう写像を決めた後、総合評価式を構成するひとつの評価式である $C(m, s)_f$ が極小になる点を検出してパラメータを決定した。しかし、こうした二段回処理の代わりに、状況によっては単に総合評価式の最小値が最小になるようにパラメータを決めても効果的である。その場合、例えば $\alpha E_0 + \beta E_1$ を総合評価式とし、 $\alpha + \beta = 1$ なる拘束条件を設けて各評価式を平等に扱うなどの措置を講じてもよい。パラメータの自動決定の本質は、エネルギーが最小になるようにパラメータを決めていく点にあるからである。

(3) 前提技術では各解像度レベルで4種類の特異点に関する4種類の副画像を生成した。しかし、当然4種類のうち1、2、3種類を選択的に用いてもよい。例えば、画像中に明るい点がひとつだけ存在する状態であれば、極大点に関する $f(m, 3)$ だけで階層画像を生成しても相応の効果が得られるはずである。その場合、同一レベルで異なる副写像は不要になるため、 s に関する計算量が減る効果がある。

(4) 本前提技術では特異点フィルタによってレベルがひとつ進むと画素が $1/4$ になった。例えば 3×3 で1ブロックとし、その中で特異点を探す構成も可能であり、その場合、レベルがひとつ進むと画素は $1/9$ になる。

(5) 始点画像と終点画像がカラーの場合、それらをまず白黒画像に変換し、写像を計算する。その結果求められた写像を用いて始点のカラー画像に変換する。それ以外の方法として、RGBの各成分について副写像を計算してもよい。

【0108】

〔画像生成に関する実施の形態〕

この実施の形態は、前提技術によってキーフレームである I 1 と I 2 の間に求められた対応点情報を格納する対応点ファイル F を利用する。対応点ファイルは第 1 のキーフレーム I 1 と第 2 のキーフレーム I 2 の中間フレームを生成するために利用される。前提技術で述べたごとく、対応点どうしの位置を補間することにより、任意の時間的位置の中間フレームを生成できる。したがって、第 1 のキーフレーム I 1、第 2 のキーフレーム I 2、および対応点ファイル F を保存しておくことで、2 つの画像間のモーフィングや滑らかな動画を生成することができ、動画の圧縮効果が得られる。前提技術でも、相当高い圧縮率で非常に美しい動画が生成されており、実験室レベルで、すでに M P E G をしのぐ成果が確認され始めている。

【 0 1 0 9 】

対応点情報は画素単位で得られるが、その場合、対応点ファイルのサイズが一般に大きくなる。そこで、第 1 のキーフレーム I 1 上にメッシュを設け、メッシュの格子点に当たる画素についてのみ対応点情報を残す工夫をしてもよい。その場合、格子点以外の点は、それを取り囲む位置にある格子点に関する対応点情報をもとに、補間によって対応点情報を生成すればよい。

【 0 1 1 0 】

図 1 8 は、実施の形態に係る画像生成装置 1 0 の構成を示す。通信部 1 2 は、ネットワーク経由で第 1 および第 2 のキーフレーム I 1、I 2 と、対応点ファイル F を受信する。受信されたこれらのデータは第 1 保存部 1 8 へ格納される。またこれらのデータが受信されたとき、その旨が性能特定部 1 4 へ通知される。

【 0 1 1 1 】

性能特定部 1 4 は、本装置の画像処理能力を検出する。性能を特定するための情報として、本装置に実装されたグラフィックアクセラレータ、またはグラフィックボードの有無、本装置の C P U のタイプおよびその動作周波数、または場合により、本装置に標準装備されているグラフィックチップの種類があげられる。これらの情報は、本装置の O S（オペレーティングシステム）から取得することもでき、本装置のシステムレジスタを読み出すことによって検出することもできる

性能特定部 1 4 はさらに別の方法として、本装置の中間フレーム生成部 2 2（後述）による実験的な中間フレームの生成をとおして本装置の画像処理能力を測定してもよい。この場合、たとえばネットワークから送られてくるデータに、あらかじめ非常に低い解像度の実験用画像データおよびその対応点ファイルを盛り込み、性能特定部 1 4 がこれを図示しない経路を介して中間フレーム生成部 2 2 に渡すことにより、実験を行うことができる。中間フレーム生成部 2 2 によって中間フレームを生成する際、その所要時間が性能特定部 1 4 に取得され、本装置の性能を実測することができる。

性能特定部 1 4 は中間フレーム生成部 2 2 による実験以外の方法で本装置の性能を検出してもよい。たとえば、コンピュータの性能を測定するために準備された、いわゆるベンチマークテストをおこなってもよい。いずれにせよ、特定された性能に基づき、本装置の処理能力が後述のランク A、B、・・・X へ分類される。

【0 1 1 2】

解像度決定部 1 6 は、性能特定部 1 4 によって分類されたランクに基づき、本装置で再生すべき動画の解像度を決定する。このために解像度決定部 1 6 は、対応テーブル 3 0 を参照する。対応テーブル 3 0 には、あらかじめ性能のランクと解像度の関係が記述されている。決定された解像度は解像度変換部 2 0 へ通知される。

【0 1 1 3】

解像度変換部 2 0 は、第 1 保存部 1 8 から第 1 および第 2 のキーフレーム I 1、I 2 を読み出し、これらを解像度決定部 1 6 から通知された解像度にしたがい、必要に応じてスケールダウン、すなわちより低い解像度への変換をおこなう。ただし、当初の解像度のままでよい場合は、第 1 および第 2 のキーフレーム I 1、I 2 はそのまま中間フレーム生成部 2 2 へ出力される。なお、スケールダウンの方法として、たとえば縦横両方向について奇数番目の画素のみを取り出すことにより、1/4 の解像度の画像を生成することができる。その他、スケールダウンには既知の任意の方法を用いてもよいが、処理全体のリアルタイム性を考慮し

、比較的高速な変換が可能な方法が望ましい。

【 0 1 1 4 】

一方、対応点ファイル F も、第 1 保存部 1 8 から読み出され解像度変換部 2 0 へ投入される。解像度変換部 2 0 において、第 1 および第 2 のキーフレーム I 1、I 2 にスケールダウン処理が施された場合、同様に対応点ファイル F の内容にも変更が施される。一例として、例えば縦横 1 / 2 のスケールダウンがおこなわれた場合、対応点ファイルに記述された画素の番号も全体に 1 / 2 に変換すればよい。解像度変換部 2 0 で必要な変換処理を施された対応点ファイルは、同様に中間フレーム生成部 2 2 へ出力される。

【 0 1 1 5 】

中間フレーム生成部 2 2 は、入力された第 1 および第 2 のキーフレーム I 1、I 2 および対応点ファイル F を利用し、前提技術で説明した方法により、補間計算によって任意の枚数の中間フレームを生成する。中間フレームの枚数は、動画の場合、1 秒間に 2 4 枚、または 3 0 枚といった数字である。こうして生成された中間フレームは、2 枚のキーフレームとともに表示制御部 2 8 へ出力される。表示制御部 2 8 は、表示装置で必要なデータフォーマットへの変換を施したのち、これらのフレームを表示順に並び替え、表示装置へ出力する。このことにより、動画の再生が実現する。

【 0 1 1 6 】

解像度のスケールダウンによって、リアルタイムの中間フレームの生成および表示が可能な場合は以上の処理でよいが、本装置の性能が非常に低い場合、いかに低い解像度へ変換してもリアルタイムの処理が不可能と判断できる場合もある。そうした場合、解像度決定部 1 6 はその旨をデータ変換部 2 4 へ通知する。データ変換部 2 4 は中間フレーム生成部 2 2 から出力されたキーフレームおよび中間フレームをバッファし、これをムービーファイルへ変換して第 2 保存部 2 6 へ格納する。ムービーファイルの例として M P E G フォーマット、A V I フォーマットなどがある。すなわち、この方法では、中間フレーム生成部 2 2 によるリアルタイムのフレーム生成は問わず、より遅い速度で生成されたフレームをいったんムービーファイルへはき出し、一連の作業が終わったのち、これを第 2 保存部

26から読み出して表示するものである。この場合、ムービーファイルの生成に所定の時間がかかるが、いったん生成されたムービーファイルは、以降切れ目のない再生が可能になる。データ変換部24は、ムービーファイルを生成する場合、その旨をユーザに通知する構成をもってもよい。

【0117】

図19は対応テーブル30の内部構成を示す。同図のごとく、対応テーブル30は、グラフィックアクセラレータの存在を示す欄100、CPUの種類を示す欄102、CPUの動作周波数（クロック）を示す欄104、リアルタイム再生が可能な最大解像度を示す欄106、および性能特定部14によって分類された性能のランクを示す欄108を含む。グラフィックアクセラレータの有無を示す欄100において「1」はアクセラレータが存在することを示し、「0」はアクセラレータが存在しないことを示す。

【0118】

同図のごとく、グラフィックアクセラレータが存在する場合、CPUの種類や動作周波数に関係なく、本装置は最高性能のランクに相当するランクAに分類されている。同様に、グラフィックアクセラレータが存在しない場合であっても、CPUの名称が「○○○」で、その動作周波数CLが750MHz以上の場合、およびCPUのタイプが「△△△」で、その動作周波数CLが600MHz以上の場合、それぞれ本装置はランクAに分類されている。これは一般にグラフィックアクセラレータが存在する場合、そのアクセラレータがCPUに代わって画像処理を担当するため、CPUのタイプや動作周波数に関係なく最高性能が期待できること、およびグラフィックアクセラレータが存在しない場合でも、CPUのタイプおよび動作周波数によっては、最高性能が期待できることに対応する。

【0119】

一方、グラフィックアクセラレータが存在せず、CPUが「○○○」でその動作周波数が400～750MHzの場合、第2ランクであるランクBに分類される。同様にCPUが「△△△」でその動作周波数が250～600MHzの場合、同様に、ランクBに分類される。ランクAにおいて、許される解像度の最大値は1200×1200である。同様にランクBにおいて許される最大解像度

は600×600である。

【0120】

一方、グラフィックアクセラレータが存在せず、しかもCPUのクロック周波数が低い場合、たとえばCPUが「○○○」でその動作周波数CLが80MHz以下の場合、同様にCPUが「△△△」でその動作周波数CLが60MHz以下の場合にはランクXに分類され、このランクはリアルタイムの再生が不可能と判断される。したがって、このランクに分類された場合、データ変換部24によるムービーファイルへの変換がおこなわれる。同様に、性能特定部14が識別できない「その他」のCPUの場合、その動作周波数に関係なく、リアルタイム処理をあきらめてムービーファイルを生成すべく、ランクXに分類されている。

【0121】

図20は、実施の形態に係る画像生成システム200の構成を示す。この構成は、いわゆるサーバ・クライアントシステムであり、サーバ50は、画像保存部52および通信部54を含む。またクライアントである画像生成装置10は、同様に通信部12および第1保存部18を含む。両者の通信部54、12はインターネット56を介して通信する。サーバ50の画像保存部52は、キーフレームおよび対応点ファイルに関するデータを保存している。一方クライアントである画像生成装置10は図18で説明した構成である。

【0122】

以上の構成による動作を説明する。まずユーザが、画像生成装置10からサーバ50へ動画の配信を要求する。サーバ50は、画像保存部52から必要なキーフレームおよび対応点ファイルFを読み出し、通信部54およびインターネット56を介して、これらを画像生成装置10へ送信する。画像生成装置10はそれらのデータを取得し、図18に示す構成にしたがって、必要な解像度の変換を加えたのちこれを表示装置に表示する。ただし解像度のスケールダウンをおこなってもリアルタイム処理が不可能と判断された場合、動画のリアルタイムな再生は断念され、生成された中間フレームがキーフレームとともにムービーファイルへ変換される。変換後のムービーファイルは、以降ユーザの望む任意のタイミングで通常のムービーとして再生される。

【0123】

以上、本実施の形態によれば、センターであるサーバ50は、ユーザ端末であるクライアントの性能に関係なく、常に同一のキーフレームおよび対応点ファイルを送信することができる。一方クライアントであるユーザ端末は、その端末内で自己の画像処理を能力を判定するため、リアルタイム再生ができる範囲で最大の解像度、すなわち最大の画質で動画を表示することができる。したがって、サーバ50はクライアント毎に対応する負荷から解放され、送信すべきデータは常に同じキーフレームおよび対応点ファイルでよく、これらの構成による動画の再生が動画の高い圧縮効果を生むことは前提技術からもあきらかである。すなわち本実施の形態によれば、伝送路のバンドをさして要求することなく、クライアント毎に異なる性能に応じた最大の画質を実現し、実用面ですぐれたシステムを提供することができる。

【0124】

以上、本発明をいくつかの実施の形態をもとに説明した。これらの機能または処理は、いずれもコンピュータプログラムの形で提供することができる。また、以上の処理の任意の組合せや組み替えもまた、本発明の態様として有効である。以下そうした変形例をいくつか挙げる。

【0125】

実施の形態では、性能特定部14がOSその他の手段を通じて装置の性能を検出した。しかしながら、性能の特定をユーザの指定に応じて実施してもよい。たとえば、ユーザはリアルタイム再生が可能であっても、これをムービーファイルへいったん記録してのちに再生を望む場合もある。その場合、ユーザがそうした選択を明示的に指示できるよう、性能特定部14はGUI（グラフィカルユーザインタフェース）を通じて、ユーザの指定、たとえば「リアルタイムで高画質」「リアルタイムで小さな画像」「ムービーファイルへの記録」などのうち希望するものを選択できる画面を生成してもよい。

【0126】

本実施の形態では、ネットワークからキーフレームおよび対応点ファイルを取得する構成としたが、画像生成装置10自らが前提技術で説明した画像マッチン

グの機能を備えてもよい。その場合、画像生成装置 10 は、2 つのキーフレームに対し、それぞれ特異点を検出する機能と、それらの特異点をもとに両者のマッチングを検出する機能とを備えればよい。またその際、前提技術でも説明したとおり、特異点の抽出によってキーフレームを階層化画像へ変換し、その解像度のもっとも粗い画像から順にマッチングを特定していく機能を有してもよい。

【0 1 2 7】

本実施の形態では、画像生成装置 10 に内蔵されるグラフィックアクセラレータの有無、またはその CPU のタイプなどが判定された。しかしながら、本装置の外部に置かれた画像処理アクセラレータを備える構成であってもよい。その場合、性能特定部 14 は、そうした外付けのアクセラレータの存否を検出しその検出結果に基づいて性能をランク分けする構成であってもよい。

【図面の簡単な説明】

【図 1】 図 1 (a) とは図 1 (b) は、ふたりの人物の顔に平均化フィルタを施して得られる画像、図 1 (c) と図 1 (d) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 0)$ の画像、図 1 (e) と図 1 (f) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 1)$ の画像、図 1 (g) と図 1 (h) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 2)$ の画像、図 1 (i) と図 1 (j) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 3)$ の画像をそれぞれディスプレイ上に表示した中間調画像の写真である。

【図 2】 図 2 (R) はもとの四辺形を示す図、図 2 (A)、図 2 (B)、図 2 (C)、図 2 (D)、図 2 (E) はそれぞれ相続四辺形を示す図である。

【図 3】 始点画像と終点画像の関係、および第 m レベルと第 $m-1$ レベルの関係を相続四辺形を用いて示す図である。

【図 4】 パラメータ η とエネルギー C_f の関係を示す図である。

【図 5】 図 5 (a)、図 5 (b) は、ある点に関する写像が全単射条件を満たすか否かを外積計算から求める様子を示す図である。

【図 6】 前提技術の全体手順を示すフローチャートである。

【図 7】 図 6 の S 1 の詳細を示すフローチャートである。

【図 8】 図 7 の S 1 0 の詳細を示すフローチャートである。

【図 9】 第 m レベルの画像の一部と、第 $m - 1$ レベルの画像の一部の対応関係を示す図である。

【図 1 0】 前提技術で生成された始点階層画像を示す図である。

【図 1 1】 図 6 の S 2 に進む前に、マッチング評価の準備の手順を示す図である。

【図 1 2】 図 6 の S 2 の詳細を示すフローチャートである。

【図 1 3】 第 0 レベルにおいて副写像を決定する様子を示す図である。

【図 1 4】 第 1 レベルにおいて副写像を決定する様子を示す図である。

【図 1 5】 図 1 2 の S 2 1 の詳細を示すフローチャートである。

【図 1 6】 ある $f(m, s)$ について λ を変えながら求められた $f(m, s)$ ($\lambda = i \Delta \lambda$) に対応するエネルギー $C(m, s)_f$ の挙動を示す図である。

【図 1 7】 η を変えながら求められた $f(n)$ ($\eta = i \Delta \eta$) ($i = 0, 1, \dots$) に対応するエネルギー $C(n)_f$ の挙動を示す図である。

【図 1 8】 実施の形態に係る画像生成装置の構成図である。

【図 1 9】 画像生成装置の対応テーブルの構成図である。

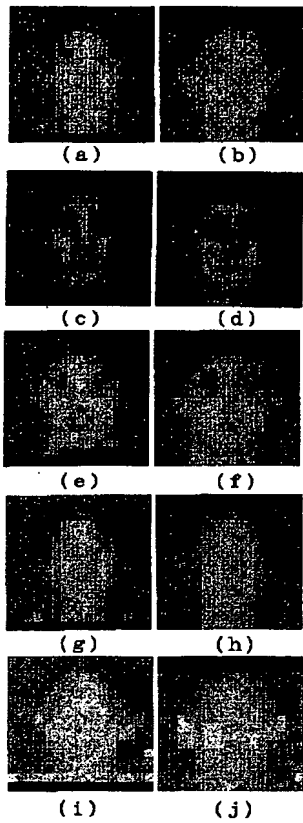
【図 2 0】 実施の形態に係る画像生成システムの構成図である。

【符号の説明】

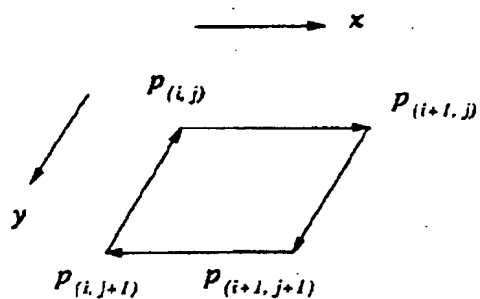
- 1 0 画像生成装置
- 1 2 通信部
- 1 4 性能特定部
- 1 6 解像度決定部
- 1 8 第 1 保存部
- 2 0 解像度変換部
- 2 2 中間フレーム生成部
- 2 4 データ変換部
- 2 0 0 画像生成システム

【書類名】 図面

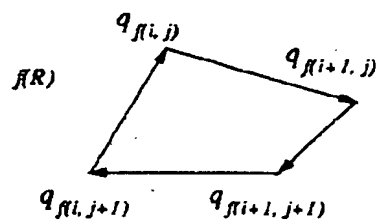
【図 1】



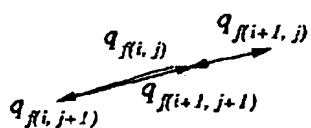
【图 2】



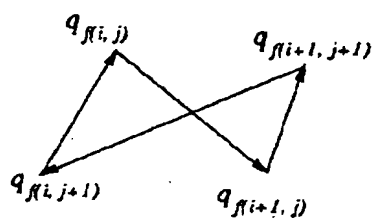
(R)



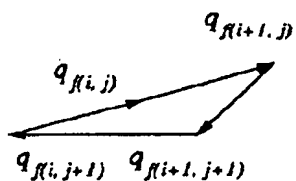
(A)



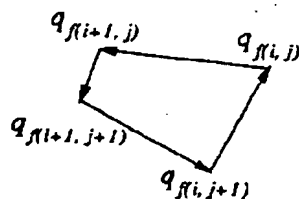
(E)



(B)

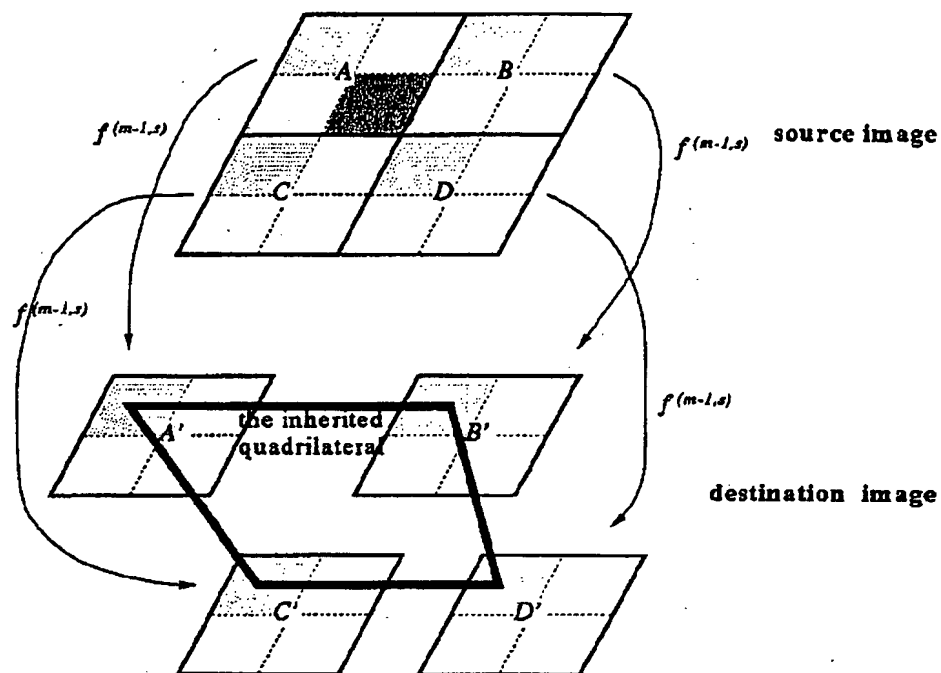


(D)

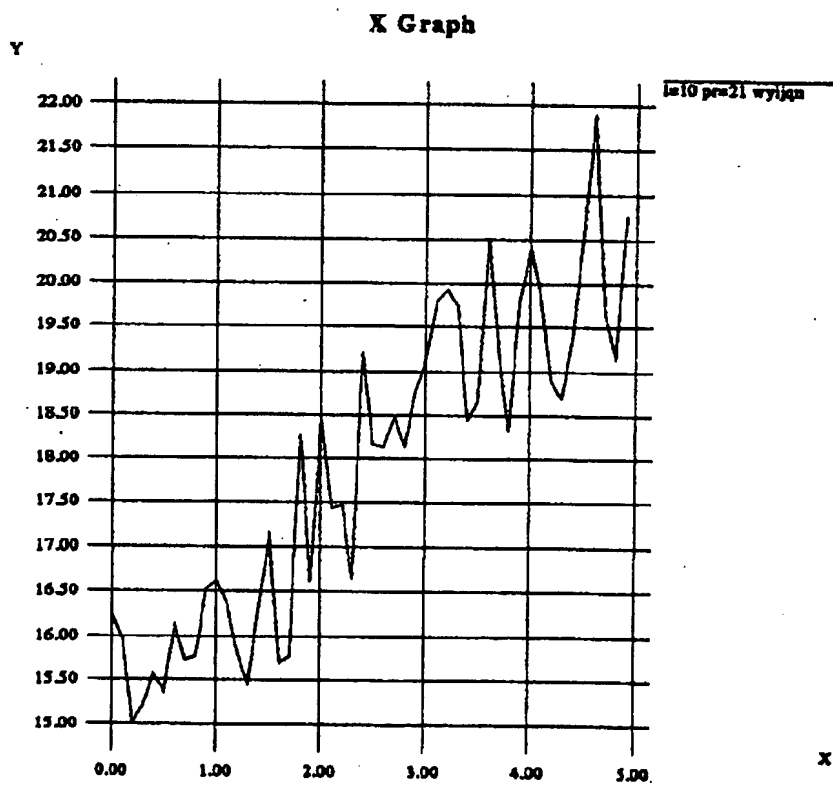


(C)

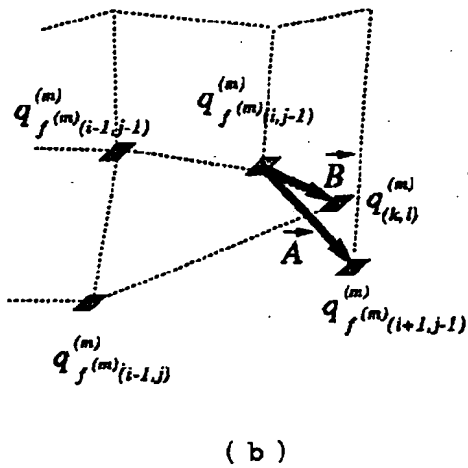
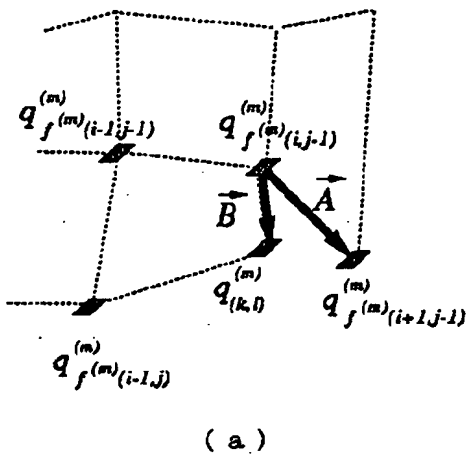
【図 3】



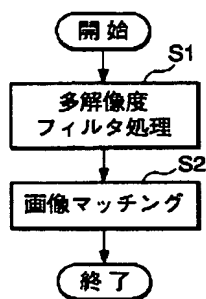
【図 4】



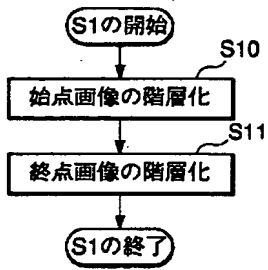
【図 5】



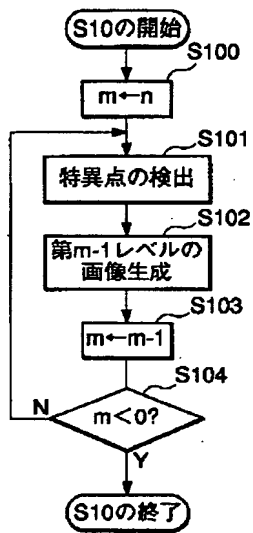
【図 6】



【図 7】



【図 8】



【図 9】

10	8		
6	3		

$p(m, s)$

3		
---	--	--

$p(m-1, 0)$

8		
---	--	--

$p(m-1, 1)$

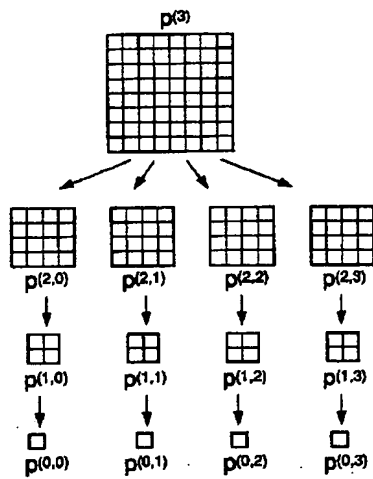
6		
---	--	--

$p(m-1, 2)$

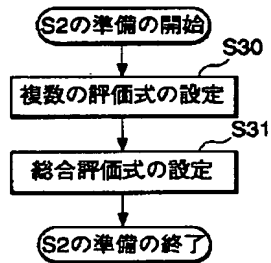
10		
----	--	--

$p(m-1, 3)$

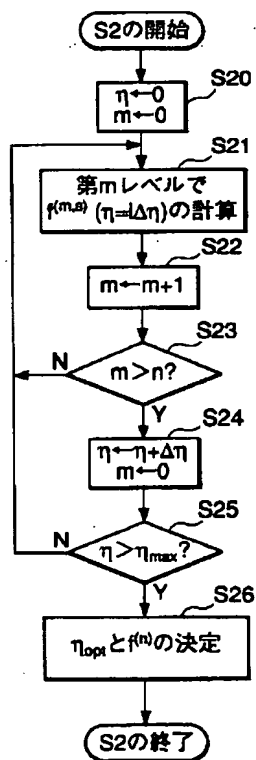
【図 10】



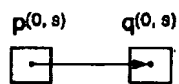
【図 11】



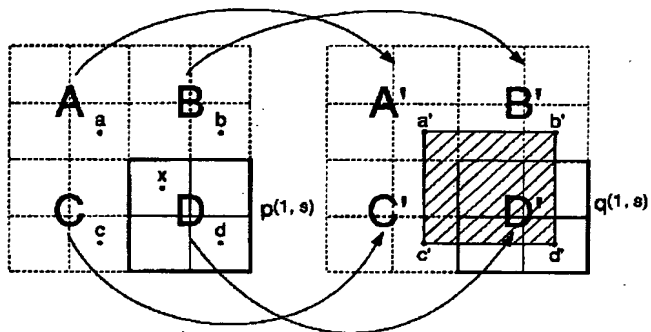
【図 12】



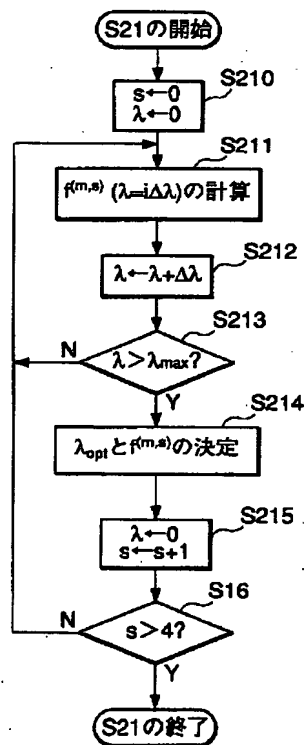
【図 13】



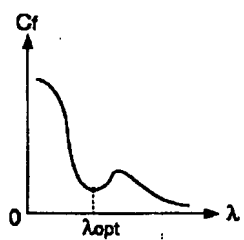
【図 14】



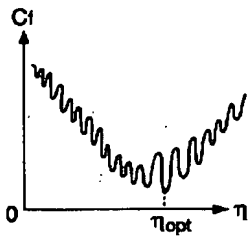
【図 15】



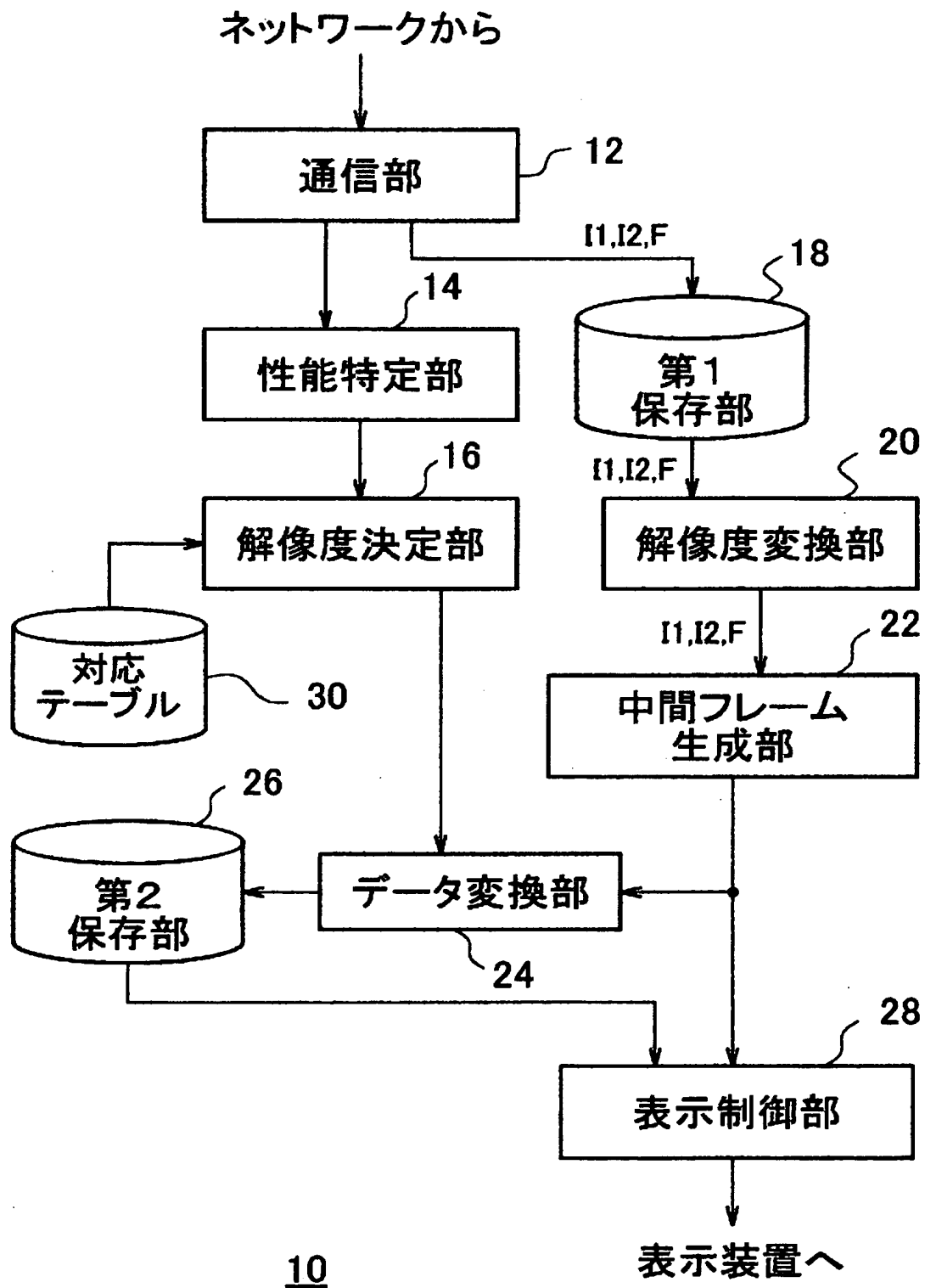
【図 16】



【图 1 7】



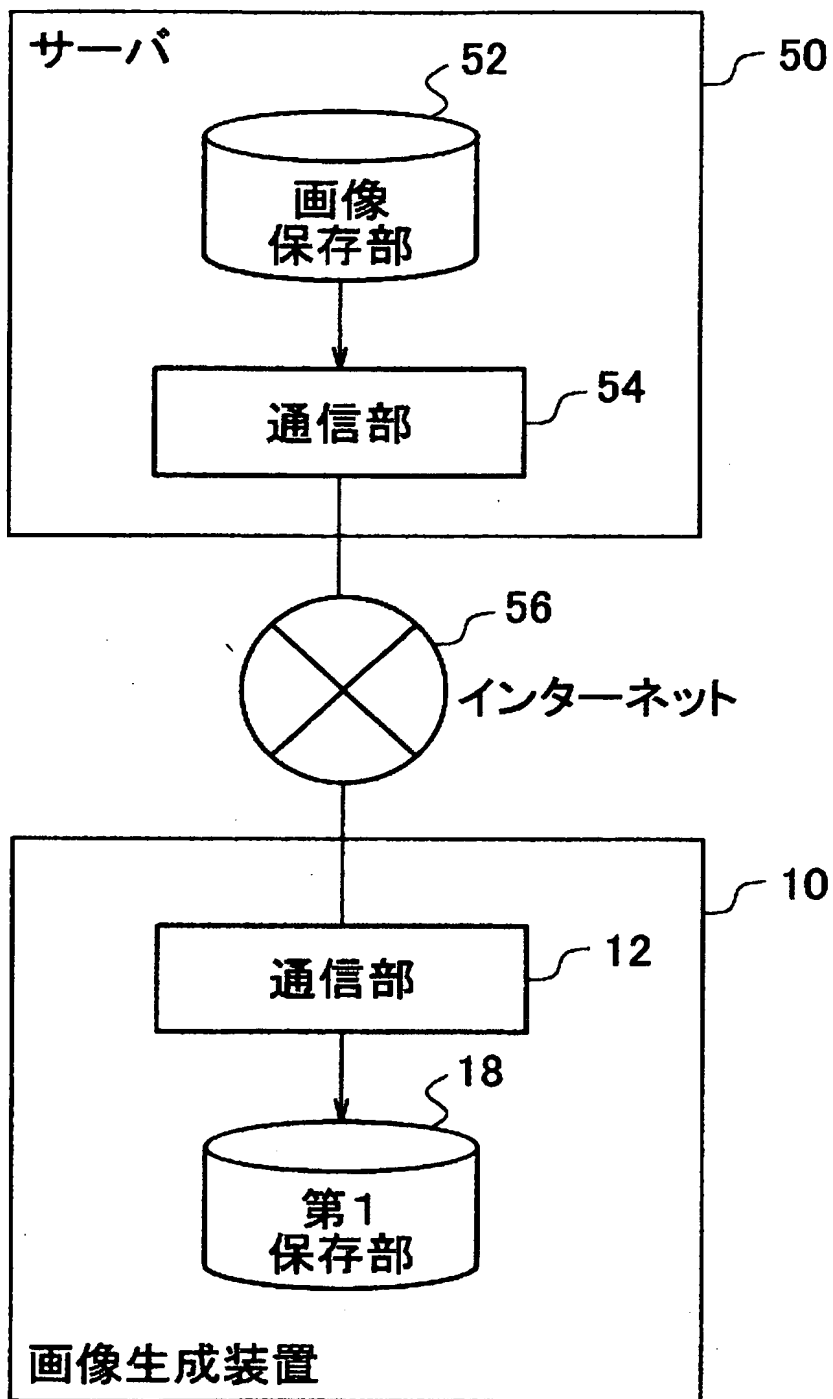
【図18】



【図 19】

100 グラフィック アクセラレータ	102 CPU	104 クロック	106 解像度	108 ランク
1	—	—	$r \leq 1200 \times 1200$	A
0	○○○	$750 \leq CL$	$r \leq 1200 \times 1200$	
	△△△	$600 \leq CL$	$r \leq 1200 \times 1200$	
	○○○	$400 \leq CL \leq 750$	$r \leq 600 \times 600$	B
	△△△	$250 \leq CL \leq 600$	$r \leq 600 \times 600$	
	～	～	～	
	○○○	$CL \leq 80$	ムービーファイル	X
	△△△	$CL \leq 60$	ムービーファイル	
	その他	—	ムービーファイル	

【図 2 0】



200

【書類名】 要約書

【要約】

【課題】 ユーザ側の端末にて、高い圧縮効果を実現する動画を高画質で表示するための画像処理に対する要望は強まっていく。

【解決手段】 第 1 保存部 1 8 は、第 1 および第 2 のキーフレーム I 1、I 2 およびそれらの対応点ファイル F を保存する。性能特定部 1 4 は、ユーザ端末の画像処理性能を検出する。解像度決定部 1 6 は、検出の結果に応じて再生すべき動画の解像度を決定する。中間フレーム生成部 2 2 は、その解像度に従い、対応点ファイル F をもとに、ふたつのキーフレームから補間によって中間フレームを生成する。

【選択図】 図 1 8

出 願 人 履 歴 情 報

識別番号 [396001980]

1. 変更年月日 1996年 1月26日

[変更理由] 新規登録

住 所 東京都港区麻布十番1丁目7番3号

氏 名 株式会社モノリス